QC
670
C2
ch.15

Cambridge
ysical Series

MODERN ELECTRICAL THEORY

SUPPLEMENTARY CHAPTERS

CHAPTER XV

SERIES SPECTRA

LIBRARIES
UNIVERSITY OF MAINE
AT ORONO

RAYMOND H. FOGLER LIBRARY
ORONO

CAMBRIDGE PHYSICAL SERIES

SERIES SPECTRA

CAMBRIDGE UNIVERSITY PRESS
C. F. CLAY, MANAGER
LONDON: FETTER LANE, E.C. 4

LONDON: H. K. LEWIS AND CO., LTD.,
136 Gower Street, W.C. 1
LONDON: WHELDON & WESLEY, LTD.,
28 Essex Street, Strand, W.C. 2
NEW YORK: THE MACMILLAN CO.
BOMBAY
CALCUTTA } MACMILLAN AND CO., LTD.
MADRAS
TORONTO: THE MACMILLAN CO. OF
CANADA, LTD.
TOKYO: MARUZEN-KABUSHIKI-KAISHA

ALL RIGHTS RESERVED

MODERN ELECTRICAL THEORY

SUPPLEMENTARY CHAPTERS

CHAPTER XV

SERIES SPECTRA

BY

NORMAN ROBERT CAMPBELL, Sc.D.

A MEMBER OF THE STAFF OF THE RESEARCH LABORATORIES
OF THE GENERAL ELECTRIC COMPANY, LIMITED, LONDON

CAMBRIDGE
AT THE UNIVERSITY PRESS
1921

Digitized by the Internet Archive
in 2026

PRINTED IN GREAT BRITAIN.

https://archive.org/details/modernelectrical0000norm_i9m7

PREFACE

THIS is the first of a series of monographs intended to supplement, and eventually to replace, my book on *Modern Electrical Theory*. Progress in the science is so rapid that it is impossible to make the book always justify its title by issuing new editions. It is proposed therefore to issue as soon as possible a series of monographs, similar to this, dealing with the most important advances in physics since the last edition and to continue the series whenever the need arises. When the present edition of the book is exhausted, those parts of it which can still be described as 'modern' will be re-written and will form other monographs of the series.

The extent of each monograph corresponds roughly to that of a chapter of the book. Accordingly, in order that there may be continuity between the book and its supplements, the monographs are called chapters and numbered consecutively with the chapters of the book. Chapters I—XIV are in the book; subsequent chapters are monographs. References from one part of the work to another are made by chapter and page numbers.

Although the form of the book, and probably much of its authorship, will be changed, its purpose will remain unaltered. It is still intended, not for experts, but for students who, having taken the usual examination courses, wish to get into touch with research. It aims at giving them such knowledge of the chief regions in which research is active that they may proceed at once to the original memoirs; much more attention is therefore given to the clear exposition of the main ideas which inspire the research than to detailed statements of the results that have been obtained from it. References to original literature are not given; for, if the book attains its object, the literature to which it is most important that the reader should turn is that which is not yet published.

The choice of Series Spectra for the subject of the first monograph needs no justification. The work which is based on Bohr's theory is undoubtedly the most important advance in pure physics since 1913. That theory will characterise an era as distinct as those

which followed the electromagnetic theory of light or the electron theory. It is already influencing so wide a range of science that it has been difficult to compress this monograph within the limits assigned to it; and the difficulty is increased because the immense accumulation of facts correlated by it are not given in ordinary text-books and not thoroughly familiar to the students for whom the work is intended. However, since the facts concerning optical spectra, other than those of hydrogen and helium, are not yet brought completely within the theory, it was decided that a brief mention of them would suffice. Again, the bearing of the theory on the constitution of atoms and molecules is reserved for a later monograph. And for the facts concerning X-ray spectra, other books, notably that of the Braggs, are available.

It will be obvious that in writing the monograph much help has been obtained from Sommerfeld's *Atombau und Spectrallinien*; though, for the sake of variety, I have diverged as widely as possible from that admirable treatise in all matters of detailed treatment.

N. R. C.

August, 1921.

CONTENTS

PAGE

I. PRELIMINARY 1

Optical spectra. X-ray spectra.

II. THE REGULARITY OF SPECTRAL SERIES 6

Wave-length and frequency. Series spectra. Lines and terms. The Principle of Combination.

III. THE ORIGIN OF HOMOGENEOUS RADIATION . . . 15

The first assumption. Homogeneous radiation and the wave-theory. Radiation and mechanics. Normal and ionised states. The ionisation potential. Absorption and emission. Carriers of the spectrum.

IV. THE STATES OF AN ATOM 31

The numerical relations of terms. Terms in optical spectra. Terms in X-ray spectra. Fundamental principles of explanation. The orbits of an electron. The selection of possible states. The mass of the nucleus. X-ray spectra. Other optical spectra. Bohr's Theory and mechanics.

V. LINES AND COMPONENTS 60

The generalisation of Bohr's assumption. Two simple examples. Elliptical orbits. Total and partial quantum numbers. Resolution into components. Conditionally periodic orbits. The Zeeman effect. The Stark effect. The fine structure. The structure of X-rays. Lines and components.

VI. THE INTENSITY OF SPECTRAL LINES 93

Intensity and probability. The Principle of Correspondence. Extension of the Principle.

VII. BAND SPECTRA 105

Regularities in Band Spectra. The rotation of the molecule.

INDEX 110

SERIES SPECTRA

I. PRELIMINARY

Optical spectra. Every element in the gaseous state possesses a characteristic spectrum, which appears under low dispersion as a finite number of bright lines that are very narrow compared to the dark spaces between them. Under higher dispersion, the lines are usually found to be complex and to consist of a number of components, of which the breadth is comparable with the distance between them; if some of the components are very close to and faint compared with others, they are called 'satellites' of the brighter components. The distinction between a line, a component, and a satellite is merely one of degree, but it is practically useful; the word 'line' will often be used to include all the components and satellites which are much nearer to each other than they are to the rest of the spectrum; the wave-length of such a complex line is the wave-length which would be measured with a dispersion too small to separate the components.

If the dispersion is continually increased, a stage is always reached at which further increase reveals no further components. But in this ultimate stage, the components are always found to have a finite breadth and to cover some small range of wave-lengths. This breadth is, however, not characteristic of the individual atoms that are the actual emitters of the spectrum. It is due to two main causes. First, the atoms are always in motion with different velocities and in different directions relative to the observer; in virtue of the Doppler effect, the frequency of the light from them received by the observer is changed in different degrees. Second, the atoms are differently situated with respect to the forces exerted by their neighbours, and these forces change the frequency slightly from that which would be emitted if the atoms were very distant from each other; the extent of this change increases with the density of the emitting substance, whereas the Doppler change increases with its

temperature. There is every reason to believe that, if the atoms were at rest and infinitely distant from each other, the lines of their spectra would be perfectly sharp; the light of each line or component would be perfectly homogeneous, that is to say, every part of it would be treated in exactly the same manner in passing through any optical system. (Cf. Chap. IX, p. 190.) This assumption, that the spectrum of an element can be resolved ultimately into perfectly homogeneous components, is fundamental in the theory we are about to discuss.

Spectra are generally excited either by feeding the element into a flame or by passing through it some form of electric discharge-arc, or spark, or the discharge at low pressure. They can also be excited as a fluorescent emission under the action of incident light of suitable frequency; very remarkable and complicated spectra of iodine and of sodium can be obtained by passing through the vapours of these elements homogeneous light of particular frequencies. The spectrum emitted is, partly but not wholly, independent of the means adopted for excitation. The collection of lines excited does not vary with small changes in the conditions of excitation, though the relative intensities of the resulting lines may change considerably. But it is possible, with many elements if not with all, to excite several perfectly different spectra by sufficiently different methods of excitation; in particular the spark spectrum of many elements is quite different from that emitted from the flame or arc; the flame and spark spectra of the same element may not contain a single line in common. This variety of spectra does not indicate any looseness of connection between the lines of the same spectrum; if any lines of a spectrum are emitted, then, in general, all lines of that spectrum will be emitted; the change from one spectrum to another is a complete change from one set of regularly associated lines to another set regularly associated. The different spectra of the same element are related as are the spectra of different elements; and any theory of spectra must interpret them as characteristic of atoms differing from each other in much the same way as the atoms of different elements. Compounds do not usually emit line spectra; they either emit continuous or band spectra, or emit the line spectra of the elements of which they are composed.

In addition to one or more characteristic line spectra, some elements, e.g. nitrogen, possess band spectra[1]. But band spectra in general are characteristic of compounds, not of elements; and it is now practically certain that the band spectrum of an element is that of its molecule, not of its atom. Throughout the greater part of our discussion we shall leave band spectra entirely out of account, and consider them only in the final section when the ideas based on the study of series line spectra have been elaborated.

Besides their emission spectra, elements possess absorption spectra; out of a continuous spectrum they absorb light in lines as narrow and as discrete as the lines of their emission spectra; these absorption lines appear as dark lines on the continuous background. It was one of the earliest discovered facts in the science of spectroscopy that the absorption and emission spectra were closely connected; indeed it was long believed that the two spectra were identical—a belief largely founded on the observation that a sodium flame will absorb the *D*-lines which it emits. But subsequent inquiry has shown that the connection does not amount to identity; the example of the *D*-lines in the sodium flame is by no means typical; for instance hydrogen in its normal condition does not absorb any of the *H*-lines that are so prominent in its emission spectrum. The relation between the two spectra will concern us hereafter, and will provide some of the firmest ground on which to build our theory.

X-ray spectra. 'Light' in the foregoing paragraphs does not necessarily mean visible light. The use of the word to include infra-red and ultra-violet radiation is familiar, but nowadays its meaning must be extended to include X-rays. Our knowledge of X-ray spectra is almost wholly subsequent to the writing of the earlier chapters of this book, and it cannot be assumed that the reader is as familiar with their most general features as he may be

[1] It is difficult to state precisely the distinction between a line and a band spectrum, for the latter also consists of discrete lines, although their separation is so much less that relatively high dispersion is needed to make it evident. The general appearance of the two classes at once distinguishes them to the eye; a scientific distinction will doubtless be made ultimately by the numerical relations between the frequencies of the lines. Cf. pp. 105—109. But while all line spectra have not been resolved into series it is impossible to lay down the criterion definitely.

expected to be with those of optical spectra. The relevant facts about them will be introduced when they become important for our theoretical discussions; but a few general remarks may be made here.

Our knowledge of X-ray spectra is based wholly on the fundamental discoveries mentioned on pp. 297—299 of Chap. XI; the wave-lengths of X-rays are measured, in accordance with the principles sketched there, by the use of crystal gratings, of which the structure is known. The method has been developed in the last few years so far that it is now possible to measure X-ray wave-lengths with an error of about 1 part in 100,000, which is nearly the accuracy attainable in optical spectra by the use of ruled diffraction gratings. Much greater accuracy (limited only by the breadth of the lines and sometimes reaching 1 part in ten million) can be attained in optical measurements by means of 'interference' methods. But these methods all depend on the specular reflection of light; and since there is no specular reflection of X-rays, the methods are not applicable to them. So far as we can see at present, measurements of X-ray spectra must always be less accurate than those of optical spectra.

The results of the measurements show that, as was suspected, hard X-rays correspond to short waves, soft rays to long waves. Many of the general statements concerning rays of different hardness made in Chap. XI can be translated directly into statements concerning rays of different wave-length. The wave-lengths of X-rays that are accessible to experiment are limited by the experimental difficulties mentioned on p. 263; they lie between 12 Å. and 0·1 Å. (1 Å. $= 10^{-8}$ cm.). The shortest waves that have been observed by 'optical' methods (of course they are not visible) are about 100 Å. in length.

The characteristic rays of the various elements, discussed in Chap. XI, are now known not to be really homogeneous. The K-rays, which seemed in absorption experiments to be all of the same quality, are resolved by the X-ray spectrograph into a line spectrum; and the same is true of the L-rays, and of the more recently discovered M-rays, which are of greater wave-length and softer than the L-rays. (It is now almost certain that there are no 'J-rays,' harder than the K-rays.) The radiation in each line of the spectrum is homogeneous to the same extent as that of the lines

of optical spectra; experimentally the lines always show a certain breadth, but there again is good reason to believe that this breadth is due to the actions mentioned before. It is with these line spectra, strictly analogous to optical spectra, that we shall be concerned hereafter. The K-rays, L-rays and M-rays of the element are not to be regarded as different spectra in the sense of p. 2; although the softer rays can be excited without the harder, the harder are always accompanied by the softer; the relation between the three spectra for any one element is quite as close as that between different parts of the same optical spectrum. There are no X-ray spectra peculiar to compounds; a chemical compound always emits the rays characteristic of the elements of which it is composed. When the characteristic radiation is excited by bombarding a target with cathode rays, it is always accompanied by some general radiation of all wave-lengths, forming a continuous spectrum.

It is also possible now to investigate more closely the characteristic absorption of X-rays. It was known previously that an element absorbed rays just harder than its characteristic radiation more than rays just softer. If a continuous spectrum is examined after passing through a plate of some element, the emergent radiation will show (besides the characteristic emission spectrum of the element) a sharp boundary, dividing a region of less from a region of greater absorption. The position of this boundary, or 'absorption limit,' is characteristic of the element. There is one such limit connected with the K-emission spectrum and situated at a wave-length just less than that of the shortest K-line. There are three such limits connected with the L-emission spectrum and probably more than three with the M-emission spectrum; but only one of these limits has a wave-length shorter than *all* the lines of the emission spectrum. However there is still an intimate connection between the absorption limit and the emission lines. For no fluorescent radiation (see Chap. XI) is emitted unless the incident radiation is of wave-length as short as some absorption limit; and when any absorption limit is reached, the fluorescent radiation emitted consists only of lines of wave-length greater than that of the absorption limit. The significance of this fact will appear presently. The collection of absorption limits may be termed the absorption spectrum.

II. THE REGULARITY OF SPECTRAL SERIES

Wave-length and frequency. The regularity of the connection between the lines of the same spectrum suggests that it may be possible to find regular numerical relations between the wave-lengths of those lines; the fundamental physical connection between them may be expected to show itself in simple numerical laws. But the search for such laws does not seem to have been undertaken seriously for at least 30 years after characteristic line spectra were discovered, and it was only in the closing years of the nineteenth century that any notable success in the search was obtained.

The finding of the laws was undoubtedly hindered by the method universally adopted for expressing the results of measurement of spectra. The lines are always described in experimental work by their wave-lengths. The wave-length varies with the refractive index, and therefore with the density and composition, of the medium in which the measurements are made; and for accurate work the necessity for reduction to some standard medium has long been recognised. The standard medium for the expression of experimental measurements is dry air at N.T.P.; but for theoretical work it is clearly better to choose a vacuum. The refractive index of air for visible light is about 1·0003; so that λ_{air} will differ from λ (the value for a vacuum) by only about three parts in 10,000. Since, however, the error of all but the very roughest measurements is much less than this, the difference is important; and the reader must remember it if he compares values given here (which are all based on λ) with values given in the ordinary tables (which are all λ_{air}). But the simplicity of the numerical relations between the lines of a spectrum does not become clear until the use of the wave-length is abandoned altogether, and in its place the frequency, ν, is adopted. If c is the velocity of light in vacuo (3×10^{10} cm./sec.),

$$\nu = \frac{c}{\lambda} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1).$$

We do not know c nearly as accurately as we know λ; moreover the values of ν calculated from λ by means of (1) are inconveniently

large (always greater than 10^{13}). Accordingly the practice has arisen of adopting as the frequency, in the expression of measurements, the quantity

$$\nu = \frac{1}{\lambda} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

in place of $\nu = \frac{c}{\lambda}$. The substitution is equivalent to adopting as the unit of time, in place of the second, the time that light takes to travel 1 cm. in vacuo; and a mere change of unit never changes the form of a numerical law. ν, given by (2), is also called the 'wave-number,' because it is the number of waves occupying a length of 1 cm. in vacuo. Henceforward, in stating all numerical results, the 'experimental frequency' ν, given by (2), and not that given by (1), will be adopted; and it will be simply termed 'the frequency.' But since, in theoretical equations, it would be inconvenient to change the unit of time, in all theoretical work the 'true' frequency, given by (1), will be used. Whenever the symbols of an equation are translated into numerals, there must be substituted for the symbol ν the experimental frequency multiplied by c, or $c\nu$ must be substituted for ν before numerical values are introduced. The procedure sounds cumbrous, but since it is that which is always adopted nowadays, it would be inconvenient to depart from it; and it will not be found actually to be as bad as it sounds. No further reference to this matter will be made.

Series spectra. When it was realised that simple relations should be sought between the frequencies of lines in a spectrum, and not between their wave-lengths, progress was rapid. But there still remain many spectra which cannot be reduced to any law. Thus, the spectrum of iron is perfectly definite, so much so that it is often used as a standard for estimating the frequencies of lines in other spectra; but the lines in it still seem distributed at random. On the other hand, all the line spectra that have been reduced completely to any order have all been reduced to an order of the same kind; if we can find any law that describes with complete accuracy the frequencies of the lines in a spectrum, that law is always of the same type. The spectra that are subject to a law of this type are now called series spectra, for reasons that will be apparent

when we discuss the law. Such series spectra are the main subject of this chapter. This limitation of the theory with which we are concerned to a particular class of spectra might seem at first sight to detract from its value; for since we have no reason to believe that the processes involved in the emission of one line spectrum are essentially different from those involved in another, it might be urged that a theory which accounts for one cannot be really satisfactory unless it accounts for all.

But this criticism is not well-founded. In the first place, series spectra form a definite group related to the other properties of the elements of which they are characteristic. Most of the elements which possess true series spectra belong to the first three columns of the periodic table; that is to say, they belong to the chemical groups of alkalis, alkaline earths, or earths. The exceptions (and we shall see that the first two are supremely important) are hydrogen, helium, and perhaps neon. Most of the elements in this class are known to possess also spectra of a different type, which have not so far been reduced to series spectra; but no element outside this class possesses a series spectrum. Again, such regularities as are found in spectra other than series spectra are quite consistent with the laws of series spectra. It is therefore possible that they may ultimately be resolved into series spectra, but that, on account of their far greater complexity, the resolution has not so far been effected. And if this explanation is correct, the reason why series spectra are not found on the right of the periodic table is apparent. For it is a fact, concordant with the theory that we are about to discuss, that spectra, even when they are series spectra, increase in complexity as we pass from the left side of the table to the right. It should be added that all X-ray spectra are series spectra.

And now we may proceed to consider what are the features common to the laws of all series spectra. There are two such features. Each has an entirely different significance for the theory, and it will be well to separate them completely in our discussion. The feature that will be considered first is not that which was discovered first or, probably, that which is best known to the reader; it is taken first because the part of the theory which explains it is logically prior to that which explains the second feature, which will not be mentioned at all until the discussion of the first is completed.

Lines and terms. This feature may be described most easily with the help of an example; the series spectrum of hydrogen will be chosen for the purpose, both because it is the simplest and because it is the most important. It should be mentioned that hydrogen possesses another and quite different characteristic spectrum, known as the 'secondary' spectrum, which has not been reduced to series; it almost certainly belongs to the hydrogen molecule. When the spectrum of hydrogen is mentioned in what follows, it will always mean the series spectrum, characteristic of the hydrogen atom, unless the contrary is stated.

The frequencies of lines in the spectrum of hydrogen are shown in the upper line of Fig. 1. The lines are arranged in three well-marked groups, that in the visible spectrum being the 'Balmer series' (so called because the numerical relations between the

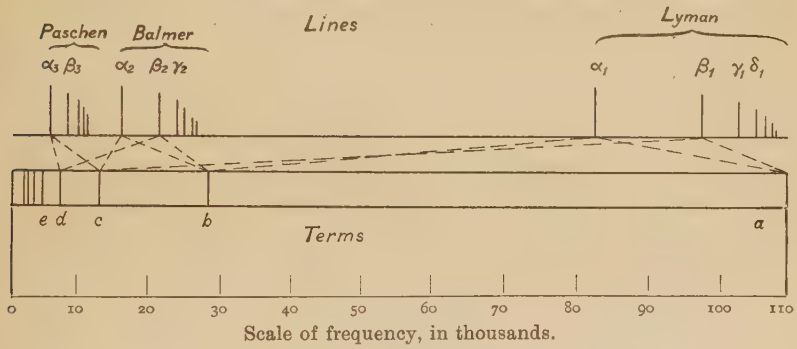

Fig. 1.

frequencies in it were first pointed out by J. J. Balmer (1885)) which contains the well-known red and blue lines, H_α, H_β, H_γ; the group in the infra-red was discovered by Paschen, that in the ultra-violet by Lyman. In each group the lines become more crowded as the frequency increases and at the same time the intensity of lines (indicated roughly by their length in the figure) decreases. The relative intensity in a single group varies somewhat with the manner of excitation; between the lines of different groups it varies so greatly with the excitation that no general statement can be made. This fact alone indicates that the groups really are distinct and that the relation between the lines in one group is closer than that of lines in different groups.

There are also many other facts pointing in the same direction. In all the properties in respect of which lines may differ (e.g. their sharpness or diffuseness, their mode of resolution in a magnetic or electric field), the lines in a single group resemble each other more than do lines in different groups. The groups form physically distinct parts of the spectrum; they are known as 'series'; and it is one of the characteristics of the kind of spectra that we are considering that they are made up of these physically distinct series[1].

On the lower line of Fig. 1 are marked certain frequencies which will be called 'terms.' These terms are related to the lines immediately above them by the following proposition, which is all-important for our discussion:

The frequency of every line is the difference between two terms; and (with certain exceptions to be explained later) every difference between a pair of terms is the frequency of a line.

Thus, if the frequencies of the lines and the terms are denoted by the letters attached to them in the figure, we have

$$\alpha_1 = a - b; \quad \beta_1 = a - c; \quad \gamma_1 = a - d; \quad \delta_1 = a - e; \quad \text{.....(Lyman)}$$

$$\alpha_2 = b - c; \quad \beta_2 = b - d; \quad \gamma_2 = b - e; \quad \text{.....(Balmer)}$$

$$\alpha_3 = c - d; \quad \beta_3 = c - e. \quad \text{.....(Paschen)}$$

(A few of these relations are indicated in the figure by dotted lines between the line of the spectrum and the pair of terms of which it represents the difference.) The Lyman series in the ultra-violet is obtained by subtracting from the first term on the right the terms to the left of it; the Balmer series by subtracting from the second term on the right the terms to the left of it; the Paschen series in the infra-red by subtracting from the third term on the right the terms to the left of it. If we proceeded in this manner to subtract the remaining terms from the fourth, fifth, ... terms on the right, we should obtain lines yet further in the infra-red, which we cannot hope to observe. When we had completed the process we should have taken the difference between every pair of terms and should find that every line is a difference between terms and every difference of terms is a line, which will be observed if it lies in a region accessible to experiment.

[1] As a matter of fact, the distinction between the different series is much less clear in hydrogen than in other elements.

But how are the terms found? Actually they are calculated from the lines; they are chosen so that the relations which have just been stated are fulfilled. And this calculation is no mere juggling with numbers, which might be carried out whatever were the facts; the fact that we can choose terms in this manner is supremely significant. Of course, however the lines lay, it would be possible to express them as differences between terms. But if the procedure were artificial, we should not find that, when we had selected the terms to account for the lines, a line would correspond to *every* difference between a pair of terms. We should have to choose a number of terms at least as great as the number of lines; and the number of differences would be greater than the number of lines; actually we find that the number of terms necessary is less than that of the lines; to account for seven lines in each group, 21 in all, we need only ten terms. Again, we should not find, as we actually do find, that the lines in a single group such as the Balmer series, which are closely related physically, were also related closely in the differences of the terms representing them.

The Principle of Combination. There is then no doubt that this method of representing lines as the difference of terms corresponds to something physically important. It was discovered by W. Ritz in 1908 and called by him the Principle of Combination. He did not express it quite in the form that has been given, or indeed introduce the conception of 'terms' at all; he expressed it solely by means of the differences between lines. It is unnecessary now to consider his form, but it will be well to notice the way in which he was led to it by consideration of the lines alone. The relations which suggested the Principle of Combination are these. (1) The frequency of every line can be represented as the difference between the frequencies of two other lines. Thus, since α_1 is $a - b$ and β_1 is $a - c$, α_2, which is $b - c$, is also $\alpha_1 - \beta_1$. But it should be observed that not every difference between two lines corresponds to a third line; thus there is no line corresponding to $\alpha_1 - \alpha_3$, for the frequency of such a line, expressed by terms, is $a - b - c + d$, which is not the difference between a pair of terms. (2) It is possible to find many pairs of lines which have the same frequency difference. Thus the following pairs of lines have the frequency

difference $a-b$:—$\beta_1-\alpha_2$, $\gamma_1-\beta_2$, $\delta_1-\gamma_2$, ... and the following have the frequency difference $a-b-c+d$:—$\alpha_1-\alpha_3$, $\beta_1-\beta_3$, ... Whenever either of the relations (1) or (2) is found in any spectrum, we suspect nowadays that it may be resolvable into terms, although it is not always possible at present to complete the analysis.

The Principle of Combination, according to which lines are the differences between terms, is the first characteristic of the class of series spectra. The class includes all the spectra of which the lines usually employed in the identification of elements form a part; e.g. the *D*-lines of sodium, the violet line of potassium, the yellow and green lines of mercury and so on. It includes indeed all the spectra which are experimentally most important, except, perhaps, the iron spectrum so often used as a standard. These other series spectra differ from the example that has been taken in the complexity of their terms. In the hydrogen spectrum these terms, as may be seen from Fig. 1, form a single simply related group converging regularly towards the left. If the terms of other optical spectra were displayed in the same manner, the eye might not detect at first any regularity in their arrangement; though, as we shall see later, there is quite a simple regularity. This complexity makes it more difficult to discover what the terms are, for the process of finding them cannot be reduced to any rule, but must depend largely upon trial and error; it is rendered easier by a fact that has been indicated already. The lines of this class of spectra can always be resolved into groups, called 'series,' which are such that all the lines of one group resemble each other more closely than lines of different groups; thus, they will be all sharp or all diffuse, all resolved by magnetic or electric fields in the same way, and their relative intensities will vary less with the mode of excitation. Such series of lines, distinguished by their physical characteristics, are always found to possess some common property when their frequencies are expressed as differences between terms; usually they all have one term in common, and the frequencies of the lines of the series are obtained by deducting the frequencies of other terms from this term. Thus the Balmer, Lyman and Paschen series in Fig. 1 are series in this sense; but they are not typical series; for though the lines of one of these series resemble each other as closely as those of a typical series, the lines of different series do not differ as greatly.

The presence of these series, distinguished from each other by the physical characteristics of the lines in them, was one of the earliest results of the detailed study of spectra; and it has given rise to the name 'series spectra.' But it is not the characteristic which is most important for modern theory, although it receives a simple interpretation from that theory. For our present purpose it would be more convenient and illuminating to call the class that we are considering 'combination' spectra; but since that term is not in common use, it will not be adopted.

If X-ray spectra had been known and studied before optical spectra, it is likely that the Principle of Combination would have long awaited discovery. For the arrangement of the X-ray terms is quite as complicated as that of any optical spectrum, and, since the number of lines known and the accuracy with which their frequencies can be measured are much less, it would have been difficult to establish with certainty the existence of terms, between which lines were differences, unless the hint had been given by the study of optical spectra. Moreover, though the lines can always be expressed as the differences between terms, there are more exceptions to the necessary converse proposition that every difference between terms represents a line; even if the right terms had been suggested, it might have appeared that they represented nothing but a numerical coincidence. But now that optical spectra have been studied, and a theory evolved to explain them, X-ray spectra are seen to provide some of the most notable examples of the Principle of Combination and some of the most striking confirmations of the theory based on it.

A word should be added about absorption spectra, whether optical or X-ray. The lines (or absorption limits) of an absorption spectrum are related to the terms of the emission spectrum in the same way as are the emission lines; their frequencies too are the differences between terms, and the terms are the same. It would seem to follow that absorption lines must always coincide with some emission lines; but these emission lines are not always observable. For we may fail to observe a line, not only because it lies in a region where observation is impossible, but also because its intensity is too low. Thus, in Fig. 1, the intensities of the lines of each series fall off rapidly towards the right; the lines are clearly converging

to a limit there, but though there is every reason to believe that the number of lines so converging is infinite, we can only observe a few of them. Generally less than 10 can be observed and measured; in a few cases as many as 40; beyond these limits the lines become so faint and so close that they cannot be distinguished. Some of the most important absorption lines lie (especially in X-ray spectra) at, or very near to, the limit to which the emission lines of a series are converging; and hence they do not coincide with any emission line that can be accurately measured and distinguished. But so long as these frequencies are represented accurately by the differences between the frequencies of the terms that are necessary to account for emission lines that can be observed, the failure to observe an emission line coinciding with the absorption line throws no doubt upon the general relation.

III. THE ORIGIN OF HOMOGENEOUS RADIATION

The first assumption. The explanation of the Principle of Combination was put forward by Niels Bohr in 1913. The paper in which his theory was expounded is one of the most remarkable in the history of science; not only for the youth of its author (who was then working as a research student under Sir E. Rutherford), but also for the completeness of the ideas contained in it. Most theories of far-reaching importance have been attained by successive approximation: this theory sprang into existence fully developed. The second part of it, to which we shall proceed later, underwent some change and development; but the first part, with which we are now concerned, stands as it did when first propounded.

Bohr's fundamental idea was that the terms of a spectrum represent and correspond to states of the system emitting the spectrum. In what follows we shall speak of this system as 'the atom,' though the possibility that series spectra may be emitted by molecules, or even larger complexes, is not completely excluded. The radiation of which the frequency is represented by $a - b$ is emitted during a transition of the atom between the states corresponding to the terms a and b.

The manner in which the terms represent and correspond to atomic states is suggested by the quantum theory of radiation (see Chap. X), which is expressed formally by the equation

$$W = h\nu \qquad \text{.....(3)}$$

In that equation W is an amount of energy, ν the frequency of some radiation, h a universal constant. In its original and crudest form this equation was interpreted to mean that radiant energy of frequency ν was always radiated or absorbed in bundles or quanta $nW = nh\nu$, where n is some integer, usually 1. It appears now that a more general and less precise interpretation must be placed on it; but of its fundamental importance and wide application in the theory of radiation there is every day more evidence. The relation

of (3) to Bohr's fundamental idea is obvious. Every atomic state must be characterised by a certain energy; the state represented by the term a will be characterised by the energy W_a, and so on. If the doctrine of the conservation of energy applies to the emission of radiation, the atom in passing from a to b must part with energy $(W_a - W_b)$. But this amount of energy, emitted during the transition, is associated by (3) with the frequency $\frac{(W_a - W_b)}{h}$; it is natural to suppose that all the energy lost in the transition appears as energy of radiation, and that the frequency $\frac{(W_a - W_b)}{h}$ is the frequency of this radiation. This frequency then will be the difference of W_a/h and W_b/h; if W_a/h and W_b/h are the terms a and b, we understand at once why the frequency of the line is the difference of the terms.

Accordingly our theory is this. The terms, calculated from the lines and now denoted by ν_a, ν_b, etc., are connected with the energies characteristic of certain states of the atom by the equations

$$h\nu_a = W_a, \quad h\nu_b = W_b \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4).$$

Homogeneous radiation, appearing as a line of the spectrum, is emitted when the atom changes from one of these states to another; and the frequency of this radiation is $\nu_a - \nu_b$.

Homogeneous radiation and the wave-theory. Before we proceed to consider the direct experimental evidence for this theory, it will be well to remark how completely at variance it is with the views of the mechanism of radiation that were prevalent before it was introduced.

According to Maxwell's theory, light is emitted by the vibrations of charged particles, and the frequency of the light emitted is connected closely with that of the vibrations. It is a difficult question (discussed in Chap. IX) whether according to this theory the light emitted actually 'consists' of light of the different frequencies into which it is resolved in the spectroscope, but this at least was agreed, that no completely homogeneous and monochromatic radiation could be emitted except from a system maintained for an infinite period in a steady state. This proposition applies to all kinds of vibrating systems, emitting all kinds of

radiation. Thus the sound vibrations from a tuning-fork are not really homogeneous, because the amplitude of the vibrations gradually decays; even if the fork is maintained electrically, some departure from homogeneity again arises, because the vibrations have been started at a finite time before the observations were made. According to this older theory, truly homogeneous radiation is an ideal which may be very nearly approached, but never attained. But Bohr's theory holds that the radiation emitted by an atom in passing between two steady states is wholly homogeneous and monochromatic; any breadth which the line is actually found to have must have been introduced after emission (e.g. by the Doppler effect); and this completely homogeneous radiation is emitted by a system which, on the older view, could least well emit it, namely one passing discontinuously between two states. Here is the clearest opposition between the new ideas and the old.

But, it may be asked, in rejecting so entirely the older view, are we not rejecting also the wave-theory of light; for the principles on which the older view of homogeneous radiation is based seem inextricably involved in that theory? And if we reject it, what is to be put in its place to explain the vast mass of laws which are unintelligible without it? The answer is this: The new theory rejects the wave-theory of light only in those applications of it for which there is no experimental evidence. The experimental evidence for the wave-theory is based on the phenomena of interference and diffraction. Now all text-books explain that in interference there is no loss of energy; the light which is absent from the dark bands is accumulated in the light bands. All the evidence for the wave-theory is based on observations in such conditions that there is no change in the amount of radiant energy of each frequency. On the other hand the theory that we are considering is applicable only to conditions in which there is change of radiant energy; energy is either emitted or absorbed. In such phenomena the wave-theory invariably suggests wrong results, of which many were discussed in Chaps. IX—XI; a striking example is found in the emission of electrons under the action of X-rays, where the electrons appear, immediately the rays are turned on, with energies which, according to the wave-theory, they could not possibly acquire in a time less than many hours. (Chap. IX, p. 289.)

In considering the properties of radiation, whether it is the emission of spectra or any other phenomenon, we must draw the sharpest possible line between circumstances in which there is a change of radiant energy and circumstances in which there is no change. To the latter the classical wave-theory applies in all its strictness; radiation, so long as its energy is not changing, behaves like a continuous train of waves. To the former the new discontinuous theory applies; radiant energy does not appear as a continuous train of waves but as a single indivisible whole. This single whole can be described and identified by means of conceptions applicable to the wave-train which appears when we examine it without changing its energy; we can thus speak of its frequency or its wave-length. But as soon as its energy changes, these conceptions become inapplicable, and we can speak only of the energy-quantum characteristic of it, which is related to the frequency by the equation $W = h\nu$.

As yet no bridge can be established to connect these two views of radiant energy; we can give no idea why radiation which, when its energy changes, appears as a quantum W, appears, when its energy is constant, as a train of frequency W/h. But the problem of building the bridge is not necessarily insoluble, and in solving it guidance will doubtless be obtained from the second part of Bohr's theory to which we shall proceed later. The two parts have been clearly separated because the division between them corresponds accurately to the division, so important for the whole study of radiation, between the conditions of constant energy and the conditions of changing energy.

Radiation and mechanics. The second conflict between the new theory and the old concerns something even more fundamental than the wave-theory of light; it concerns the principles of classical dynamics. The new ideas involve not merely the denial of those principles, but the rejection of the very conceptions on which they are founded. For according to Bohr's theory, the states of an atom differ by finite amounts of energy; the terms are separated by finite distances, and between two consecutive terms there is no term and, therefore, no state of the atom. Now it is an absolutely fundamental assumption of classical mechanics that the possible states

of any system form a continuous series, consecutive members of which differ infinitesimally in energy or in any other property; and that the passage between two states, different in a finite degree, occurs through an infinite number of intermediate states, differing infinitesimally; so much is implied by the use of differential equations.

The two contradictory views might possibly be reconciled by describing the finite number of states allowed by Bohr's theory as the only *stable* states, leaving the remainder to be regarded as possible but unstable. But the distinction between stable and unstable states is quite foreign to Bohr's theory, and it is by no means certain that all the states corresponding to terms are actually stable according to classical mechanics. If we are going to reject classical mechanics at all—and it is impossible today to adhere to it in dealing with any of the phenomena of radiation—we had better do so completely and not allow ourselves to be hampered by the relics of a discarded tradition. Classical dynamics, perfectly valid within a certain sphere, formed a self-consistent whole; and if one part of it has to be abandoned when we pass out of that sphere, there is not the smallest reason apart from mere mental inertia for retaining any other part. At any rate, if anyone does care to think or to speak of the states intermediate between Bohr's possible states as existent but unstable, he must remember that he is introducing something which does not belong to that theory; and if from his conception of unstable but existent states he deduces consequences inconsistent with the theory, he must not think he has raised an objection to it. It may possibly be an objection that the theory is inconsistent with classical tradition—though others will hold this an advantage in a department of science in which that tradition has always failed—but it is no objection that it is inconsistent with ideas which it definitely rejects; on this score objection may be taken rather to attempts to put new wine into old bottles.

However, the recognition of unstable states, intermediate between stable states and not represented by terms, introduces new difficulties. For if, in a transition between states of energy W_a and W_b, the atom passes through unstable states of intermediate energy W_x, why is not radiation of frequency $(W_a - W_x)/h$ or

$(W_x - W_b)/h$ emitted? Of course we might simply deny that it is and leave the matter, but that is not a very satisfactory solution. Moreover this question merely indicates a much more general objection, which has actually been raised and cannot be overcome by such simple means. It is really a variant of an objection which first appears in history as Zeno's paradox of motion, and has been cropping up at intervals ever since. It is argued that something must be happening between the moment when the atom leaves state x and the moment when it arrives at state y; it is in this period that the emission of radiation occurs; Bohr's theory holds that this radiation is determined by the final state y that has not been reached as much as by the initial state x that has been left; but how can the action be determined by something that has not occurred? It may be replied that this difficulty is inherent in the classical view of continuous motion as much as in any other. There too a system passes from one state to another: how can the intervening actions, including that which determines that the system arrives at the second state and at no other, be determined by the final state? The theory of continuous motion answers the question by denying that there are any intervening actions; it supposes the successive states to be so close, infinitesimally close, that there simply is nothing at all between them. Bohr's theory (though I cannot speak for its author) makes the same answer; there simply is nothing between the two states of an atom between which a transition occurs with the emission of light; there are no intervening actions which can be determined by the initial but not the final state[1]. The atom changes from one state to another, and that is all there is to be said about it. The theory will not admit that anything true can be said about intervening states; it refuses any validity to the conception, and denies that it can enter into any physically important statement. Of course if anyone cares to

[1] It should be pointed out that light is often emitted during the transition between terms that are not consecutive. Is then the radiation corresponding to the intermediate terms emitted? Considerations of energy make it difficult to believe that it is; we must regard the intermediate terms as not intermediate in the sense that the atom has to pass through them in its transition between extremes. But this view does not seem to present any new difficulty not inherent in the conception of a discontinuous change. When a system changes *discontinuously* from a to b, there is nothing whatsoever intermediate between those two states.

say that there *is* something about which nothing significant can possibly be said, he is at liberty to do so; but he can hardly expect intelligent people to discuss the matter with him.

Normal and ionised states. Much of the evidence for the first part of Bohr's theory is derived from its success when combined with the second part. But since it has been insisted that it is important to distinguish the two parts, it will be well to consider what evidence for the first part can be adduced that is wholly independent of the numerical relations between the terms. We must ask what reason there is to believe (1) that there are discontinuously separated atomic states, (2) that the energy of these states is connected by the quantum relation with the terms of the spectrum, (3) that radiation is always the accompaniment of a change from one of these states to another. We will take these questions in order.

(1) It is known from experiments which have nothing to do with spectra that atoms can exist in at least two states that are discontinuously separated, namely the 'normal' electrically neutral state, and the positively charged or 'ionised' state in which the atom has lost an electron. Some atoms are known to be capable of losing more than one electron; and our first guess might be that the different states of the atom corresponded to different numbers of electrons lost and different positive charges. But that hypothesis is clearly impossible; thus hydrogen, which all evidence shows to be capable of losing only one electron, must have as many states as there are terms in its spectrum; the number of these terms is at least 40, and there is every reason to believe that it is really very much greater. On the other hand, the intimate connection between ionisation and the emission of the spectrum (Chap. IX, p. 205) indicates strongly that the ionised state and the normal state are two of the states represented by terms; the suggestion occurs that the other states may be intermediate between these extremes and consist of some form of partial ionisation. If there are such states, an atom in one of them should require less energy to ionise it than an atom in the normal state.

In recent years our knowledge of the ionisation of gases has been greatly advanced by experiments (in which Franck has been

especially prominent) on the interaction between neutral atoms and electrons possessing known speed and energy. In general it appears that collisions between electrons and atoms do not result in appreciable ionisation until the energy of the electrons exceeds a definite 'ionising potential'[1], characteristic of the atom. On the other hand there is a definite potential lower than this, above which the electron begins to lose energy when it collides with a molecule; in other words, an electron can give to an atom an amount of energy less than that required for ionisation. This energy is very soon lost by radiation (as we shall see presently); if its reception causes the atom to pass to another state, that state does not persist long. It is not to be expected therefore that many atoms should be in such a state at any one time, and, since it is only a very small fraction of the atoms that are ionised at all, it is not to be expected that many atoms should be ionised while in this state and while they require a lesser energy than the normal for ionisation. It is difficult therefore to detect certainly the presence of such atoms, but there are phenomena, especially in the ionisation of helium, which can be plausibly interpreted as due to the presence of partially ionised atoms, requiring less than the full ionising potential for ionisation.

Another very interesting argument, which indicates that, if the terms represent atomic states, these states are intermediate between the normal and ionised atom, is due to Sommerfeld. If the terms of a single spectrum represent states intermediate between the normal and the ionised state, in which the atom has lost an electron, and if the atom is capable of losing two or more electrons, what are the terms that represent the states intermediate between losing one electron and losing two? Surely they must be terms of some other spectrum. Now it is known that each of the alkali metals, and those of the alkaline earths and earths, possess two spectra, which are known respectively from their mode of excitation as the arc (or flame) and spark spectra. If the arc spectrum is emitted

[1] The energy of an electron is always expressed by the voltage through which it must fall in order to acquire that energy. That is to say, the energy W is replaced by V the voltage, where $W = eV$ and e is the charge on an electron ($4{\cdot}774 \times 10^{-10}$ E.S.U.). Throughout this chapter, when atomic energies are estimated numerically, it will always be by volts and not by ergs.

in transitions between states intermediate between the normal and that in which one electron is lost, the spark spectrum may be emitted in transitions between states intermediate between the loss of one and two electrons. The view is plausible, for the conditions in the spark (the greater energy of the ionising electrons and the greater intensity of ionisation) are favourable to the detachment of more than one electron from a single atom.

Further, modern theories of the structure of the atom (Chap. XVI) indicate that the electrons which are removed in ionisation (except in the inactive gases) are the 'valency' electrons, of which the alkali metals in their normal state possess one, the alkaline earths two, the earths three, while the inactive gases possess none. If then a metal of the earths loses an electron from its atom, the remainder will possess as many valency electrons as an atom of an alkaline earth metal. But its spectrum is determined by states in which the conditions of its valency electrons vary; for, if the terms of the spectrum represent different stages towards ionisation, they probably represent different conditions of the electron which is removed in ionisation. It is therefore suggested that, since the arc spectrum of an alkaline earth and the spark spectrum of an earth each represents states intermediate between the possession of two valency electrons and the possession of one, there should be considerable similarity between these spectra. In the same way, there should be a similarity between the arc spectrum of an inactive gas and the spark spectrum of an alkali metal, or between the arc spectrum of an alkali metal and the spark spectrum of an alkaline earth.

Such similarity is actually found. While the arc spectrum of an alkali metal is a relatively simple series spectrum, the spark spectrum resembles the arc spectrum of an inactive gas in being extremely complex and not yet resolved into series and terms. Again, in the arc spectrum of an alkaline earth, the lines (and therefore the terms) are all close triplets; but in the spark spectrum they are doublets, and they are doublets also in the arc spectrum of an alkali. The arc spectrum of an earth is a mixture of single lines and triplets, but their spark spectra resemble the arc spectra of the alkaline earths in consisting of triplets. The rule, suggested by our theory as interpreted by Sommerfeld, that the spark spectrum

of a group in the periodic table resembles closely the arc spectrum of the preceding group, is confirmed by observation.

The ionisation potential. (2) If we assume that the unionised and the ionised states are two of the possible states of the theory and that (except in atoms which can lose more than one electron and therefore have two spectra, which for our purpose may be regarded as those of different elements) all the other possible states are intermediate between these two, then there is definite evidence for equation (4). For the difference in energy of these two extreme states must be equal to the work which has to be done in order to make the atom pass from one to another; that is to say, it is the work which must be done to ionise the atom. The terms of the hydrogen spectrum (and the same is true of all the other spectra to which the theory applies) converge on the left towards the frequency 0 as shown in Fig. 1; on the right they cease at a definite frequency ν_a, which is the largest term permissible by the Principle of Combination. Accordingly the extreme difference of frequency is ν_a, and the extreme difference of energy should be $h\nu_a$, which should be equal to the work required to ionise the atom.

Since this work, W_0, can be measured experimentally by the observations mentioned on p. 22, the relation

$$W_0 = h\nu_a \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

can be tested. Table I gives the result of the test for all elements for which measurements are available. The first column names the element, the second gives W_0 in volts, the third ν_a in cm.$^{-1}$, the fourth the ratio of the second to the third; this ratio should be constant and equal to h.

TABLE I.

Element	W_0 (volts)	ν_a (cm.$^{-1}$)	h (calc.)
Mercury	10·38	$6{\cdot}17 \times 10^4$	$1{\cdot}257 \times 10^{-4}$
Zinc	9·5	7·58	1·25
Cadmium	8·9	7·25	1·23
Sodium	5·13	4·14	1·238
Potassium	4·1	3·50	1·17
Rubidium	4·1	3·37	1·22
Caesium	3·9	3·14	1·24
		Mean	$1{\cdot}230 \times 10^{-4}$

It will be seen that h is quite as constant as the errors in measuring W_0 permit. The value of h in the last column is expressed with the volt as the unit of energy and $1/(3 \times 10^{10})$ sec. as unit of time. To convert it to erg sec.$^{-1}$ we must multiply by $\frac{4{\cdot}774 \times 10^{-10}}{300 \times 3 \times 10^{10}} = 5{\cdot}30 \times 10^{-23}$. We thus obtain $6{\cdot}52 \times 10^{-27}$ erg sec.$^{-1}$, which accords well with the value obtained from Planck's theory (Chap. X, p. 226). The confirmation of the prediction is complete.

Other similar facts, which are often quoted in support of Bohr's theory and are relevant to our inquiry, may be noticed here. It was mentioned before that electrons with energy too low to ionise an atom are yet able to give up energy to it. This reaction always results in the emission of some line in the spectrum of the element; the energy of the electrons at which this transference of energy and emission of radiation occurs is called the 'resonance' or 'radiation' potential of the atom. It is found that the radiation potential is related to the frequency of the line emitted by the quantum relation (3). If a table is drawn up similar to Table I, in which for the ionising potential there is substituted the radiation potential, and for the extreme term ν_a the frequency of the line emitted by collision with electrons of that potential, we should obtain in the last column the same constant ratio. The concordance is quite as good and the number of instances in which measurements have been made considerably greater[1]; but the results have not been quoted as evidence for Bohr's theory, for they do not distinguish between it and any other which involves the acceptance of the quantum relation (3). For here ν is the frequency of emitted radiation, whereas in Table I it is a term; it is the physical interpretation of terms and the application to them of (3) that is really characteristic of Bohr's theory.

[1] The radiation potential is becoming a powerful method of investigating the extreme ultra-violet spectrum, where the older methods are difficult. Knowing the relation between radiation potential and frequency of the line emitted, by measuring the radiation potentials of an element (for it may have several) we can discover the position of some of the lines in its spectrum, even if they lie in a region where absorption is so universal and powerful that the radiation can only travel an appreciable distance in a high vacuum. By the use of it the present gap between optical and X-ray spectra is rapidly being bridged.

Absorption and emission. (3) Though it is impossible to distinguish sharply between the evidence for different parts of our theory, which are mutually supporting, it may now be taken as proved that there are many states of the atom intermediate between the normal and ionised states, and that the energy of these states is related to the frequency of the terms by (4). We are then led to the third question, what is the evidence for believing that the emission (or absorption) of homogeneous radiation always represents a transition between the atomic states?

In discussing this matter we must determine which term represents which state. The two extreme terms, we have decided, represent the normal and ionised states, but it is still doubtful whether in Fig. 1 a represents the normal and 0 the ionised, or vice versa. It might seem at first sight that since work has to be done on the normal atom in order to ionise it, the energy of the ionised state must be greater; and therefore that a, represented in Fig. 1 as possessing the greater frequency and the greater energy, must be the ionised state. But there is a great deal that is arbitrary in the way the terms are drawn in that figure. We are concerned only with differences of energies and with the distances between terms. These distances would have been unaltered if the whole set of terms had been shifted any distance in either direction along the axis or if the whole set had been reversed and the positions of a and 0 interchanged so that a had the lesser energy[1]. The arrangement adopted was selected partly for ease of drawing the dotted lines, but more because it is suggested by the second part of the theory; but it is not permissible to deduce from it any conclusions which would not follow if the whole set of terms were displaced in any manner or reversed.

The decision between the alternatives is provided by the study of absorption spectra; at the same time important evidence of the kind we are seeking is obtained. Since absorption involves a loss of radiant energy, it is reasonable to suppose that the absorbed light does work on the atom; if a transition between states takes

[1] Of course, if there were only the two terms discussed so far, reversal would alter nothing; but since there are intermediate terms, even if the set were reversed, a, separated from its neighbour by the distance ab, would be distinguished from 0, separated from its neighbour by a much shorter distance; it is relevant to ask which of these extreme terms, so distinguished, represents the normal state.

place during absorption, it must be in the direction from the normal to the ionised state, for it is such transitions that require the expenditure of work. But the atoms before absorption are in the normal state, and therefore the term representing the normal state must be one of those between which transition occurs. Accordingly if our theory is correct, we should expect all the absorption lines of any element to have one term in common; if our expectation is fulfilled, it is this term that must represent the normal state.

The expectation is fulfilled with complete accuracy. The absorption spectrum is always much simpler than the emission spectrum, and it is simpler because the only lines which occur are those which represent differences between one term, common to all lines, and the remainder. This is one of the most beautiful confirmations of the theory, for the prediction is so intimately connected with its fundamental ideas. In the hydrogen spectrum it is the term a that is common to all absorption lines, and the absorption lines are therefore simply the Lyman series; there are no absorption lines in the visible or infra-red regions. In other spectra it is the term which corresponds (in a sense yet to be explained) to the term a in the hydrogen spectrum, and the absorption lines coincide with the lines of the 'principal series.' The D-lines of sodium belong to this series, and they are absorption lines also.

The relation holds only so long as the absorbing atoms are all originally in the normal state; if they are being ionised or partially ionised by collisions with electrons, that is to say if they are in the condition in which they emit radiation, then some signs of absorption may occur at other emission lines, indicating the absorption of energy from the incident light by atoms already partially ionised. Thus it has been shown that hydrogen through which a powerful discharge is passing shows some feeble absorption corresponding to its lines in the visible spectrum. But since the number of atoms in any but the normal state is always very small compared with that of the normal atoms, such absorption is always very weak. It may be mentioned that this fact, but not the remainder of those discussed here, is explained equally by the older theories.

If, when energy is absorbed from the incident light, the atom passes from the normal to a partially ionised state, it is to be expected that the energy will be re-emitted as radiation when the

atom returns to the normal state. Accordingly it is found that absorption in absorption lines is always accompanied by the fluorescent excitation of emission lines[1]. The study of the fluorescent emission, especially when the incident radiation is homogeneous, provides additional evidence for the theory. For suppose that the incident radiation had the frequency of the line γ_1 in Fig. 1, which is the difference between the terms a and d. Then, since a represents the normal state, the transition associated with the absorption will be from a to d; the effect of the incident light will be simply to change some of the atoms from state a to state d. When the disturbed atoms return to the normal state, they may not all achieve the reverse transition at one step; they might return first to c, then change from c to b, and last from b to a. But they will all start from d, and accordingly in the fluorescent spectrum we should expect those lines to be especially prominent which are the differences between two terms of which one is d. Here again theory and experiment agree; the fluorescent spectrum excited by the absorption of the line $(a - x)$ consists almost entirely of lines $(x - y)$, where y is a or a term intermediate between x and a. There are many other variants of the same argument. Thus, it is easy to see that the lines associated with the radiation potentials (see p. 25) by the relation (3) should be the lines of the absorption spectrum, and all have in common the term representing the normal state. It is unnecessary to proceed further; all the predictions from such arguments are found to be confirmed by observation.

But a word should be added about the absorption limits in the X-ray spectrum. In the K-spectrum (which we may take for simplicity as typical) there is only one absorption limit, of which the frequency is greater than that of any emission line. The absorption limit, in fact, represents the difference between the two extreme terms; there are no absorption limits (as there are absorption lines in optical spectra) representing the difference between one of the extreme terms and the terms intermediate between the

[1] Since the light, absorbed from a beam travelling in one direction, is emitted in all directions, the fluorescence, even if the light emitted is of the same frequency as that absorbed, appears experimentally as an absorption of the original beam. There is always some true absorption of the light and conversion of radiant energy into other forms. But this absorption is general and not confined to sharp absorption lines. It must be admitted that the mechanism of this general absorption is obscure.

extremes. This feature, interpreted by the theory, must mean that in the absorption of X-rays a transition is possible only from the normal state to the 'ionised' state and not to intermediate states, while in emission transitions can occur between intermediate states, as well as the 'ionised' state, and the normal state. Kossel has pointed out a very attractive explanation of this difference between X-ray and optical spectra. The electron which is detached in the 'ionisation' that represents one extreme term of the X-ray spectrum is not a valency electron, but one of those that form part of the permanent structure of the atom; and the intermediate states represent a transference of this electron from one part of that structure to another and not, as in optical spectra, to a position altogether outside the atom. But in the normal state of the atom all the positions which an electron can occupy and still form part of the permanent structure are occupied; so that, if one electron is detached, it cannot find a resting-place anywhere within the atom; it must either return to its original position (when there will be no absorption) or be ejected from the atom altogether (when the frequency of the light absorbed will be that corresponding to a transition between the extreme terms). The single absorption line corresponding to this transition is thus explained. On the other hand, once the electron has been completely ejected, an electron from any part of the atom can move in to occupy its place; these transitions represent emission, and emission can therefore occur corresponding to a transition between any two states of the atom, including those intermediate between the extremes. We see also why fluorescent radiation can only be excited by incident radiation harder than any of the emission lines, and as hard as the absorption limit; for in order that an electron may be able to change from one position to another, one of the electrons must be completely ejected from the atom; this complete ejection corresponds to the high frequency limit of the emission spectrum and to the absorption limit.

Carriers of the spectrum. But finally we have to glance at another source of evidence of a totally different kind, which is not so satisfactory. If all the lines of the optical emission spectrum represent transitions of the valency electron which do not amount

to ionisation (except the single limiting line, never actually observed, which represents ionisation), then throughout the whole process of the displacement of the electron, by absorption of energy from light or colliding electrons, and its subsequent return to the normal position, the atom must be not ionised. The electron may be displaced, but it is not displaced so far that the atom with its valency electron ceases to behave as an electrically neutral body. But there are circumstances in which we can determine with some certainty whether the atoms concerned in the emission are or are not electrically neutral. One class of such circumstances is mentioned on pp. 206, 207 of Chap. IX. It is said there that the interpretation of the observations is somewhat difficult; but Stark, as a result of his extended studies on the 'moving' and 'resting' lines of the positive ray spectrum, is confident that the lines of the series spectrum of most elements are emitted from atoms which are positively charged, or ionised, at the moment of emission. (His conclusions apply especially to the principal series, which is that of the lines having in common the term corresponding to the normal state.) This fact, if it is true, is directly contrary to the predictions of Bohr's theory; and it seems difficult to suggest any possible modification of it which would overcome the discrepancy without introducing others more serious. Stark's conclusions have been challenged, and other evidence, from other forms of discharge in which a decision might be made between charged and neutral carriers of the spectrum, has been adduced both in support of them and in contradiction to them. This is not the place to discuss the complicated tangle of conflicting evidence. It may turn out that Stark's observations can be explained in some other manner than that which he advocates; or it may turn out that his conclusions are not as fatal to the theory as appears at first sight. But the matter should be borne in mind, because it seems to provide the only instance in which experimental evidence is directly contrary to the theory.

IV. THE STATES OF AN ATOM

The numerical relations of terms. The second part of Bohr's theory seeks to explain why the terms of the spectra have the values that are determined experimentally from the Principle of Combination. We shall start by stating briefly what these values are. The facts are probably familiar, in outline at least, to most readers; but mention of them has been carefully avoided so far, because it is important to realise that the Principle of Combination and the conclusions that may be based on it are wholly independent of the numerical relations between the terms. It is true, and it is important for the theory, that the principle applies most simply to spectra of which the terms are related numerically in a certain way; but it is quite conceivable that it might apply to terms related in some perfectly different way.

First let us note that any numerical relation that may be found between the terms is almost sure to be of a certain type. If there is any such relation, it must be expressed by an equation of the form

$$\nu = f(x, y, z, \ldots, a, b, c, \ldots) \quad \ldots\ldots\ldots\ldots\ldots\ldots (6),$$

where ν is the frequency of a term, $x, y, z, \ldots$ variables which change from term to term, and $a, b, c, \ldots$ constants which are the same for all the terms between which the relation holds. Since the values of ν are separated from each other by finite intervals, the values of f must be similarly separated, and f must have no value at all within these intervals. If the variables $x, y, z, \ldots$ were continuous variables to which any value might be assigned, the discontinuity of f could be attained only if f were an essentially discontinuous function. Functions of continuous variables can be invented which have only a limited number of values separated by finite intervals; but they are very complicated, artificial, and intractable analytically. It is much more natural to suppose that f is a continuous function, such as usually occurs in physics, but that $x, y, z, \ldots$ are only capable of a limited number of values. If this alternative is adopted, it is immediately suggested that $x, y, z, \ldots$ are integers, for integers form

a natural discontinuous series separated by finite intervals. From the outset then we are prepared to find that the variable terms in any equation (6), stating the numerical relations between the terms, are always integers. This expectation is dependent only on the most general considerations and is quite independent of quantum theory; on the other hand, it may be legitimately expected that a quantum theory will prove especially suitable to explain such relations.

The simplest numerical relation between terms is found in the hydrogen spectrum shown in Fig. 1. Here all the terms are given by

$$\nu = C \pm \frac{N}{m^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7),$$

where m is any integer, and N about $1{\cdot}097 \times 10^5$ cm.$^{-1}$. N is called Rydberg's constant. The arbitrary constant C and the ambiguity of sign are necessary on account of the considerations mentioned on p. 26; since we are only concerned with differences of terms, C cannot be formally determined and the ambiguity of sign cannot be formally resolved. But the facts already discussed indicate that it will probably be convenient to put $C = N$ and to take the negative sign. For distances between the terms show that the term a is that for which $m = 1$; this term represents the normal state and the transition from this state to any other requires the expenditure of work. It is therefore the state of least energy, and, since we are going to make the energy of a state proportional to the term representing it, as in (4), it is convenient to call the energy of this state and the corresponding term zero. This course will be adopted when we proceed to theoretical deduction; but in the statement of numerical results it would be much more convenient to put $C = 0$ and to choose the positive sign; and this procedure is adopted by the usual practice. Once more, as in the choice of units (p. 7), the quantities involved in theoretical equations have not quite the same meaning as those of which the numerical values are stated; but no confusion need arise.

A short comparison of experimental values with those given by (7) may be useful. In Table II the first column gives the conventional nomenclature of the lines of the Balmer series; the second the frequency of these lines; the third the factor which, multiplied

by N, should give this frequency according to (7); the fourth the value of N obtained by dividing the second column by the third. It will be seen that the values of N agree as accurately as the lines can be measured[1].

TABLE II.

	ν (cm.$^{-1}$)	$=N\times$	N
H_α	15237	$\frac{1}{2^2}-\frac{1}{3^2}=\frac{5}{36}$	109706
H_β	20570	$\frac{1}{2^2}-\frac{1}{4^2}=\frac{3}{16}$	109707
H_γ	23038	$\frac{1}{2^2}-\frac{1}{5^2}=\frac{21}{100}$	109705
H_δ	24379	$\frac{1}{2^2}-\frac{1}{6^2}=\frac{2}{9}$	109706
H_ϵ	25187	$\frac{1}{2^2}-\frac{1}{7^2}=\frac{45}{196}$	109705
		Mean	109706

Terms in optical spectra. In only one other optical spectrum are the terms given by a formula as simple as (7). This is the spark spectrum of helium, first discovered in the stars, but since obtained in the laboratory. In this spectrum, which will concern us presently, N is very nearly, but not quite, four times that for the hydrogen spectrum. In the remaining optical spectra that have been resolved into series, the terms are given very nearly by

$$\nu = C \pm \frac{N}{(m+a)^2} \quad \text{.....(8);}$$

m is again any integer, a some fraction characteristic of the terms represented by (8) and N, which is a universal constant, has the same value as for hydrogen. The expression on the right is conventionally written (m, a), so that (8) becomes

$$\nu = C \pm \frac{N}{(m+a)^2} = (m, a) \quad \text{.....(9).}$$

(8) is known as Rydberg's formula from its discoverer. It is not

[1] The finite breadth of the lines, due to the presence of several components, makes any figure after the fifth insignificant. The frequency of individual components can be measured with much greater accuracy, and, when the numerical relation of these components is known, N determined more accurately. (See p. 86.)

quite accurate, and a modification has been proposed by Ritz, which gives ν implicitly and contains another very small constant α:

$$\nu = C \pm \frac{N}{(m + a + \alpha\nu)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(10).$$

But usually for brevity we shall assume that (8) is quite accurate. A set of terms represented by (8), with varying m but constant a, is called a series of terms. (A series of terms must be distinguished from a series of lines, although the two are always closely connected.) Every spectrum of which the terms are given by (8), and not by (7), has several such series of terms, characterised by different values of the constant a. In a typical spectrum these series are divided into four groups, which are conventionally distinguished by substituting for a the letters s, p, d, b. Thus the four groups are

$$\nu_s = (m, s), \quad \nu_p = (m, p), \quad \nu_d = (m, d), \quad \nu_b = (m, b).$$

This distinction is not arbitrary, for the four groups of series are distinguished by the lowest value of m which occurs in any term of them. Thus in (m, s) m may have any value from 1 upwards; in (m, p) it may have any value from 2 upwards; in (m, d) any value from 3; in (m, b) any value from 4[1]. The normal state of the atom always corresponds to $(1, s)$.

Each of the four groups may include several series. There may be, and usually is, not one series of terms (m, p), but several series (m, p_1), (m, p_2), (m, p_3), etc., each of which fulfils the condition for

[1] This distinction might be abolished, if *any* value, not necessarily a proper fraction, were permitted for the constants p, d, b. For the term $(2, p)$ is the same thing as $(1, 1+p)$, and $(3, d)$ is the same as $(1, 2+d)$. If in place of the fractional values usually given to p, d, b, we used the values $1+p, 2+d, 3+b$, then the lowest value of m would be 1 in each case. But it is probable that the usual practice is theoretically significant and that the difference in the lowest value of m in the various kinds of term is important.

But in the term (m, s) a practice of this kind is actually adopted. The term is usually written $(m+\frac{1}{2}, s)$, and the first term $(1{\cdot}5, s)$; that is to say, the value given to s is less by $\frac{1}{2}$ than that which would be given if the term were written (m, s). The reason appears to be this: If the term is written $(m+\frac{1}{2}, s)$, then s does not differ from p, d, b more than these differ from each other; if it is written (m, s), s has to be greater by about $\frac{1}{2}$ than any of p, d, b. This fact shows that there is a real difference between the (m, s) term and the others, but it is not yet certain whether that difference is expressed better by a difference in m or by a difference in s. Since Bohr's theory does not suggest fractional values of m, the latter course has been adopted here. But the former is more general, and the reader is sure to meet it elsewhere.

a p-series, namely that m varies from 2 upwards; and similarly several other series (m, s), (m, d), (m, b). Thus in the sodium spectrum (m, p) is represented by two series (m, p_1) and (m, p_2); and the doublet characteristic of the spectrum, of which the D-lines are typical, consists of lines involving corresponding terms of these two series. In the very complicated neon spectrum which has recently been resolved into series (though it shows some anomalies), there seem to be eight s-series, ten p-series, and seven d-series. (Usually, however, there is only one s-series.) This great complexity of the terms must not be taken to indicate any uncertainty in the application of the Principle of Combination; for, numerous though the terms may be, they are very much less numerous than the lines; the complexity of the terms is almost infinitely less than that of the lines which are ordered by means of them.

These more complex spectra could never have been resolved into terms if there were nothing but numerical relations to guide us. In the hydrogen spectrum all the lines in a given part of the spectrum form a single series of lines and involve a common term, and the terms are therefore easy to find. But in other spectra the various series of lines may be intermingled in the same part of the spectrum, and there would be no clue to the analysis were it not for the differences in the physical characteristics of the lines to which attention has been drawn. Now that the analysis is complete, these physical characteristics can be related to the series of terms involved in the different series of lines. The following is a list of the series of lines recognised at present, with their names, roughly in order of importance; the importance is judged by the intensity of the lines, the accessibility of the region of the spectrum in which they occur, and the number of elements in which they have been discovered. The least value of the variable m in the second term is always greater by 1 than the constant number in the first.

Principal Series ..... $\nu = (1, s) - (m, p)$
First (or Diffuse) Subsidiary Series $\nu = (2, p) - (m, d)$
Second (or Sharp) Subsidiary Series $\nu = (2, p) - (m, s)$
Bergmann Series ..... $\nu = (3, d) - (m, b)$
Third Subsidiary Series ..... $\nu = (2, p) - (m, p)$
Diffuse Principal Series ..... $\nu = (1, s) - (m, d)$
Sharp Principal Series ..... $\nu = (1, s) - (m, s)$

It must be remembered that each series of lines may be a group of series in the more complex spectra, just as the series of terms is a group of series. The Principal Series always contains the strikingly characteristic lines of an element, such as the D-lines of sodium.

It should be added that connections are found between the values of p, s, d, b in the spectra of elements associated in the same chemical group; but the discussion of them is not necessary for our purpose.

Terms in X-ray spectra. In stating the facts concerning X-ray spectra, it is more convenient to speak of lines than to speak of terms, for, though there is no doubt that there are terms, they are not so easily studied.

X-ray spectra differ from optical spectra in the fact that the resemblance between the spectra of different elements is so close that it is possible to describe all X-ray spectra at the same time; indeed it is only by so doing that the regularities can be properly displayed. These regularities were first made evident by Moseley in a paper published in 1914; his method of stating the facts is shown in Fig. 2. In that diagram the abscissae are the atomic numbers (Z) of the elements[1], every tenth element being indicated by its usual symbol; the ordinates are the values of $\sqrt{\nu/N}$, where ν is the frequency of a line in the X-ray spectrum characteristic of the element, and N is Rydberg's constant. Only one line, the softest of the strong lines, in each of the K-, L-, M-spectra is shown. The range over which the three spectra extend in the case of the element of highest atomic number shown is indicated by the vertical line at the end of each curve. It will be seen that no element is shown with more than two spectra, namely either a K- and L-, or an L- and M-; but this limitation is mainly due to the fact that the range within which X-rays can be observed experimentally is that

[1] The atomic number is the ordinal number of the element in a series arranged in order of increasing atomic weight. Thus the atomic number of hydrogen is 1, of helium 2, of lithium 3, of beryllium 4, and so on. A departure from this order has to be made when an irregularity occurs in the periodic table; thus argon (at. wt. 40) has to be put before potassium (at. wt. 39); also it has to be assumed that three elements have still to be discovered, namely higher homologues of caesium and iodine and one of the rare earth metals. (See further, Chap. XVI.)

of the diagram. There is not the smallest doubt that all elements which have an L- and M-spectrum have also a K-spectrum; in fact among the radioactive elements, the K-spectrum can be observed among the γ-rays emitted, which are harder than any X-rays that

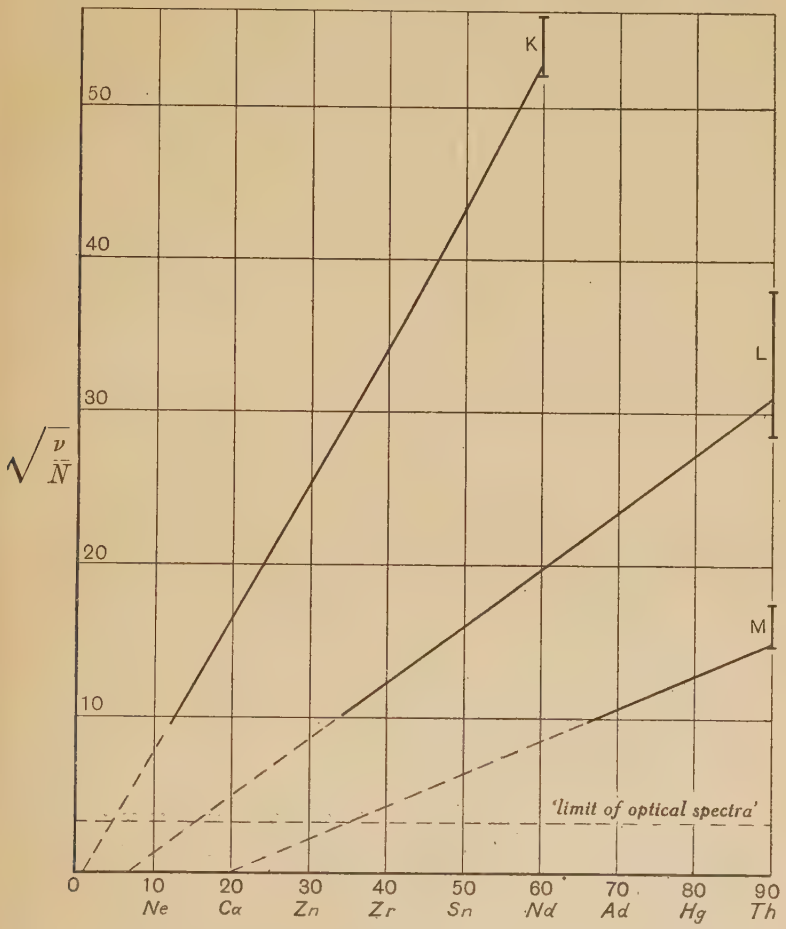

Fig. 2.

can be produced artificially. It is not so certain, for reasons that will appear later, whether every element which has a K- and L-spectrum has also an M-spectrum.

The observed points cannot be shown because they all lie within the width of the lines drawn. The variation of the frequency of

the three lines shown with the atomic number is perfectly regular; there is, for instance, no trace of irregularity in passing through the group of the rare earths ($Z = 57$ to 71), which interrupts so markedly the regularities of the chemical properties and would have been prominent throughout the curve if atomic numbers had been replaced by atomic weights. The same perfect regularity would have been seen if any other lines in the spectrum had been chosen for plotting; for in general there is exact correlation between the lines in the spectra of different elements, and to every line in the spectrum of one element corresponds a line in the spectrum of every other in such a way that if these lines are plotted against atomic numbers a smooth curve such as those shown in the diagram will be obtained. (Exceptions are found among the elements of lower atomic number which do not always possess all the lines characteristic of those of higher atomic number.) Moreover the curves are nearly straight lines. They are not quite straight, but are slightly concave upwards in the part remote from the origin, but in the earlier portions they are very nearly straight. (The concavity is more marked in the curves representing the higher frequency lines in the L-, M-spectra.) It is permissible therefore, though some uncertainty attaches to the process, to extend them to cut the horizontal axis as shown by the dotted lines; the point of intersection is always on the positive side of the axis, and its distance from the origin increases as we proceed in the direction K, L, M. All the curves representing the different lines of the same series (K, L, or M) cut the axis at very nearly the same point.

The same facts may be expressed algebraically by the equation, valid for the smaller values of Z,

$$\nu = N(Z - z)^2 . A \quad \text{.....(11).}$$

Here z is a constant characteristic of the series (K, L, or M), but the same for all lines in it and for all elements, and A a constant peculiar to the line and the series, but the same for all elements.

As has been said already, the mere experimental facts would not be sufficient to justify the application of the Principle of Combination or to determine with any accuracy what are the terms. But there are indications of the kind described on p. 11 that the Principle of Combination is applicable and that there are terms. There are six lines recognised at present in the K-spectrum,

denoted by the letters $\alpha, \alpha', \alpha_2, \alpha_3, \beta, \gamma$; fourteen in the L-spectrum, $\alpha,\alpha',\alpha'',\beta,\gamma,\delta,\epsilon,\zeta,\eta,\theta,\upsilon,\phi,\chi,\psi$, and six in the M-spectrum, $\alpha,\alpha',\beta,\gamma,\delta,\epsilon$; but new lines are still being separated, and the list is probably not complete. Among these lines the following relations, of the kind called (2) on p. 11, are found, between the frequencies of corresponding lines in all elements:

$$\left.\begin{array}{c} K_\alpha - K_{\alpha'} = L_{\alpha'} - L_\beta = L_\delta - L_\gamma = L_\eta - L_\epsilon = L_\theta - L_\zeta \\ L_\phi - L_\upsilon = L_\psi - L_\chi \\ M_\alpha - M_\beta = L_\alpha - L_{\alpha'} \end{array}\right\} \ldots(12).$$

On the other hand the following relations of the kind called (1) are nearly true, but not strictly true:

$$K_\beta - K_\alpha = L_\alpha;\ K_\alpha - K_\gamma = L_\gamma;\ K_\beta - K_\gamma = M_\alpha;\ L_\alpha - L_\gamma = M_\alpha \quad (13).$$

These are the main facts concerning numerical relations in spectra; we may now proceed to consider their explanation.

Fundamental principles of explanation. But we must first determine on what principles the explanation is to be founded. A few years ago the problem seemed perfectly definite. In order to explain the spectrum of hydrogen (or any other substance) it was necessary to devise some system of particles moving under forces, in the manner laid down by Newtonian mechanics, in periodic orbits such that, when the orbits were submitted to a Fourier analysis, the frequencies of the various terms in the Fourier series were the frequencies of the lines in the hydrogen spectrum. Further, since the electromagnetic theory of light was accepted, it was necessary that these particles should be charged and that the forces should be related to the charges and to each other in the manner laid down by Maxwell's equations. Nobody doubted what were the conceptions in terms of which the explanation must be stated, or the fundamental relations between these conceptions on which the explanation must be based. The only problem was to find exactly what relations consistent with the fundamental principles would solve the problem.

Today for reasons that have been given already these principles must be abandoned, and the theory has already adopted assumptions directly inconsistent with them. Moreover the facts that have just been stated provide fresh evidence of their inadequacy. For we

have seen that the frequencies of the lines in a single spectrum can be represented by a formula in which the term that changes from one line to another is the square of an integer. On the other hand, all frequencies in a Fourier series are necessarily related to each other by formulae in which the term that changes from one overtone to another is an integer, and not its square. Of course, since the square of an integer is also an integer, it might be possible to devise a mechanical system in which only those overtones occurred for which the corresponding integers happened to be squares; but such a system would necessarily be of enormous complexity; it could not be any natural and simple kind of system; and as a matter of fact nobody ever succeeded in inventing one.

On the other hand, the new principles that we require must cover part at least of the ground included in classical mechanics. We have abandoned the conception of radiation being related directly to the periodic vibrations of the radiating system, but we retain the conception of states of the system characterised, like those of mechanics, by the possession of a definite energy. Classical mechanics could predict, from any assumed structure of the atom, what those states must be; we hold that it must predict wrongly; but our alternative principles must predict these states and predict them rightly.

There is here undoubtedly a difficulty. Hitherto all theoretical physics has been a direct development of the work of Newton. Even Maxwell's theory, which is the most important tributary that has flowed into the main stream of thought, has much in common with Newtonian mechanics. Its characteristic vectors are closely related to Newtonian forces, and it resembles the Newtonian theory of dynamics in the form of its mathematical expression; both are expressed in differential equations of the second degree. By the end of the nineteenth century men had come to believe that all possible physical theory must be a continuation of the same development, and that a theory, if it was to be satisfactory, must conform to, or at least be consistent with, the conceptions of Newton and Maxwell. There are still some who maintain that view, and for them the newer theories are meaningless. If they mean by an explanation some theory based on Newtonian mechanics, then of course any theory not so based is not an explanation; but they must face the alternatives of admitting that some physical laws must remain

eternally inexplicable, or of admitting that explanations of other kinds are possible.

But, if we adopt the second alternative, how are we to find new principles? The answer is surely clear. We must find the new principles in the same way as we found the old. Newtonian mechanics (or its developments, such as Hamiltonian mechanics) or Maxwellian electrodynamics were not, like the Mosaic Law, imposed by divine inspiration on a passive and wondering mankind; they are merely the result of successive generalisations. The elaboration of their principles has proceeded side by side with the discovery of the laws that they have been used to explain. As new laws have been brought within their scope, the theories by which they are explained have undergone some extension, and the ideas involved in them alteration and generalisation, until now the connection between the theories and the simple laws on which they were founded originally is almost too remote for recognition. It is *not* true that the Hamiltonian equations of dynamics, applicable to any system the state of which is described in terms of generalised coordinates, state nothing that is not involved in Newton's theory; and it is *not* true that Maxwell's equations state nothing that is not involved in Ampère's theory of electromagnetism or in Faraday's theory of induction. The theories of Newton, Ampère and Faraday are merely very special cases of these more general theories; the more general theories reduce to the special cases if certain limitations are imposed on the relations between the variables involved in them. In their general form they assert far more than is involved in the special cases.

The main evidence for the truth of the theories, within the range over which they are as fully applicable as ever, is that they do explain the laws which they profess to explain. The laws, which, in the advanced state of the theories, appear as consequences of them are equally and at the same time the foundation of them and the evidence for them. No other foundation or evidence is possible. But their appeal is very greatly enhanced by the simplicity of their form. When Hamilton proposed to express all dynamics in the beautifully simple and symmetrical form of equations (41, 42), or Maxwell all electrodynamics in the equally attractive form so familiar to us today, then, even if we had not known what the theories

were about or what the symbols meant, we should have felt that a theory so beautiful had a special claim on our attention. The successive generalisations, which have extended the scope of the theories, have at the same time simplified and embellished their form.

The new principles must follow the same course of development. We start with the laws that are to be explained, but with no principles whatever on which an explanation may be based. We consider those laws first in their simplest form, and allow them to suggest principles, often with the help of analogies drawn from other branches of physics. Proceeding to the more complex laws, we shall find that the principles first suggested need amendment and extension, and we shall always seek that extension which gives the theory the widest application together with the simplest and most attractive form. If we can find principles, simple in form but wide in application, which explain adequately a great part of the known facts, and are not obviously in conflict with those of which at present no complete explanation can be given, then we shall consider that the problem has been solved. However greatly the principles at which we may arrive differ from those of the older physics, we shall hold that the kind of evidence which established the older principles is sufficient also to establish the newer; and that our principles if they are sufficiently wide and sufficiently simple are proved by their success in explanation. For, once more, no other kind of proof is conceivable.

The orbits of an electron. Let us begin then with a guess at the general structure of the atom. According to a suggestion, first indicated by Nagaoka, but now associated with the name of Rutherford who adduced strong experimental evidence for it, an atom of atomic number Z consists of Z electrons, each with a negative charge e, grouped round a small nucleus bearing a positive charge $E = -Ze$. The oppositely charged particles attract each other according to the inverse square law, and, if they were at rest, would coalesce. But they are supposed to be in relative motion, which enables them to remain apart, pursuing permanent orbits the form of which is determined by their velocities and by the forces between them. According to this view the hydrogen atom of atomic

number 1, to which we shall confine our attention for the moment, consists of a single electron with a nucleus $E = -e$. The orbits which the electron can follow are then those of a planet round the sun; they are all ellipses (including circles as special cases) of which the dimensions depend on the relative velocity of electron and nucleus, while the eccentricity depends on the direction of original projection.

Up to this point Bohr's theory coincides completely with earlier views. It assumes that the possible states of an atom are to be found among the orbits deduced from classical mechanics for electrons revolving round a nucleus. And there is no objection to the assumption; for, though ultimately a departure must be made from its principles, their complete validity within a certain range shows that the new theory must have some features in common with the old.

Consider then the simplest form of orbit, the circle, for the simplest atom, hydrogen. Let μ be the mass of the electron, infinitely small compared with that of the nucleus, a the radius of the orbit and ω the angular velocity of revolution. Equating centripetal and centrifugal forces, we have

$$\frac{eE}{a^2} = \mu a\omega^2 \quad \text{.....(14).}$$

The energy of the system is the feature important for our purpose. It was suggested on p. 32 that this energy was most conveniently reckoned from the normal state; but we do not yet know what this state is, and must start with some other zero. If we choose, as is most convenient, the state in which the electron is infinitely distant from and at rest relative to the atom, the kinetic energy of any orbit is $\frac{1}{2}\mu a^2\omega^2$, the potential energy $-\frac{eE}{a}$; the second term is negative because the system does work and loses energy when the electron and nucleus approach under their attraction. Thus

$$W = -\frac{eE}{a} + \tfrac{1}{2}\mu a^2\omega^2 \quad \text{.....(15)}$$

$$= -\frac{eE}{2a}, \text{ from (14) .....(16).}$$

If a_0 and ω_0 are the radius and angular velocity of the orbit

constituting the normal state and if we now reckon energy with this state as zero, we have

$$W = \frac{eE}{2a_0} - \frac{eE}{2a} = \tfrac{1}{2}\mu a_0^2\omega_0^2 - \tfrac{1}{2}\mu a^2\omega^2 \quad \ldots\ldots\ldots(17).$$

But from (4), $W = h\nu$, where ν is a term, and from (7), adopting the identification $C = N$,

$$W = h\nu = hN\left(1 - \frac{1}{m^2}\right) \quad \ldots\ldots\ldots(18).$$

If (17) rightly represents the energy of a possible state, (17) and (18) must be identical; since $m = 1$ for the state a_0, ω_0, this condition is fulfilled if

$$a = m^2 \cdot \frac{eE}{2hN} \quad \ldots\ldots\ldots(19),$$

$$\omega = \frac{1}{m^3} \cdot \sqrt{\frac{8h^3N^3}{e^2E^2\mu}} \quad \ldots\ldots\ldots(20).$$

(19) and (20) give the radii and angular velocities of the only circular orbits that can be possible states of the atom according to the principles that have been enunciated so far. It will be observed that the possible radii are in the ratio of the squares of the integers, the angular velocities in the inverse ratio of the cubes, but otherwise the quantities depend only on universal constants, the same for all substances. We have now to introduce some principle to explain why these radii and angular velocities alone are possible. Of course we might accept the fact as ultimate and simply lay down as a fundamental principle that only these radii are possible in the interior of an atom, but a very short inquiry would convince us that such a principle would lead us into difficulties when we attempted to apply our ideas to any spectrum but that of hydrogen. Moreover the assumption does not appear cogent and reasonable of itself; we want some wider principle and one that has a greater intrinsic probability.

The selection of possible states. The basic idea of the principle suggested by Bohr is that the method of selecting the possible states of the atom from the orbits satisfying (14) and (15) must be expressible in terms of the quantity h and of no other quantity that has not been already introduced. It must be observed that (so far

as the theory of spectra is concerned) this is a wholly new assumption which is not implied in any way by the ideas with which we have dealt hitherto. So far h has been a quantity which enters when the energy of atomic states is converted into homogeneous radiation; it is (to take a loose analogy) something like the J which enters when mechanical energy is transformed into thermal energy; it is not necessarily involved in the determination of the possible states, any more than J is involved in the determination of the possible temperature of a substance. The new assumption is that h is involved not only in the transformation of atomic into radiant energy but also in the determination of atomic energy.

h so far has been the ratio of an energy to a frequency. The atomic states are characterised by an energy and also by a frequency, namely $\frac{\omega}{2\pi}$, the frequency of revolution of the electron, which, it must be remembered, is not the frequency of the radiation emitted. It is therefore natural to inquire whether the ratio of W to $\frac{\omega}{2\pi}$ may not be simply related to h. No simple relation can be found if W is reckoned from the normal state given by (17), but if we reckon from the ionised state given by (15), we find from (20)

$$W \Big/ \frac{\omega}{2\pi} = h \,.\, m \,.\, \left(\frac{\pi^2 e^2 E^2 \mu}{2h^3 N}\right)^{\frac{1}{2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (21);$$

or, inserting the following numerical values

$$e = -E = 4{\cdot}774 \times 10^{-10},$$

$$\frac{e}{\mu} = 5{\cdot}31 \times 10^{17},$$

$$h = 6{\cdot}55 \times 10^{-27},$$

$$N = 1{\cdot}097 \times 10^{5} \times 3 \times 10^{10}$$

(for we must multiply our N by c, since in the other values the unit of time is the second), we get

$$\frac{2\pi W}{\omega h} = 0{\cdot}499\ m.$$

Since the values used are not certain to within 1 part in 500, we may identify 0·499 with $\frac{1}{2}$, and conclude that the frequency of the

terms in the hydrogen spectrum is consistent with the assumption that the possible atomic states are all such that

$$\frac{\text{energy}}{\text{frequency}} = \tfrac{1}{2}\, mh \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(22),$$

where m is an integer.

In form this principle might be plausible as a statement of the universal principle for which we are seeking, but there are difficulties in accepting it. For example, Planck's theory of complete radiation is based on the principle that the energy of an oscillator with one degree of freedom is given by $\frac{\text{energy}}{\text{frequency}} = mh$, a relation which differs from (22) by the presence in the latter of the factor $\frac{1}{2}$. Of course it might be that this factor occurs because the orbits we are considering have two degrees of freedom; we might suggest that the general principle is $\frac{\text{energy}}{\text{frequency}} = \frac{mh}{n}$, where n is the number of degrees of freedom. But the energy in Planck's relation is estimated from a different zero from that in (22). However we shall not pursue the matter further, because this is not actually the suggestion made by Bohr which has led to such notable results. It has been mentioned only to indicate the kind of principle that we are seeking and the kind of way in which we must seek it.

Bohr's suggestion arises from the observation (first made by Nicholson) that h, which was originally introduced as the ratio of an energy to a frequency, i.e. the product of an energy by a time, may also be regarded as a moment of momentum, i.e. the product of the momentum of a moving body by its distance from a point measured perpendicularly to the velocity; it has the same 'dimensions' as a moment of momentum, namely ML^2T^{-1}. Now a moment of momentum is an extremely important quantity in central orbits, i.e. those described under the action of a force directed to a fixed point. It is proved in any text-book of particle dynamics that throughout any such orbit the moment of momentum is constant[1].

[1] The moment of momentum is usually denoted in such text-books by h. Of course the agreement of this symbol with that of Planck's h is a mere coincidence, for Planck himself never seems to have thought of h as anything but the product of an energy and a time. The coincidence is not really very strange, for the number of unappropriated letters of the alphabet is limited. Here, for distinction, the moment of momentum is written p.

Now the moment of momentum in the electron orbits we are considering is

$$p = \mu a \omega \,.\, a = \mu a^2 \omega \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(23),$$

and from (19) and (20)

$$\frac{p}{h} = m \,.\, \left(\frac{e^2 E^2 \mu}{2h^3 N}\right)^{\frac{1}{2}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(24).$$

Inserting numerical values as before, we find

$$\frac{p}{h} = \frac{m}{6{\cdot}29} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(25).$$

The denominator on the right-hand side is very nearly 2π; and we may therefore write, consistently with the facts that we are examining and the ideas that have been introduced hitherto to explain them,

$$p = m \frac{h}{2\pi} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(26).$$

(26), which asserts that only those orbits are possible states of the atom for which the moment of momentum is an integral multiple of $p/2\pi$, is the original form of the characteristic principle of the second part of Bohr's theory. Perhaps it does not look very attractive at first sight, for why is the $\frac{1}{2\pi}$ there? The principle does not become really inevitable until it is developed along the lines suggested by Sommerfeld, which we shall consider later. In the meantime we shall consider some of its consequences, merely noting that all the striking consequences deduced in the following pages would result equally well from the earlier principle (22), which we have rejected largely on the score of its incompatibility with Planck's oscillator. But the principle we accept is scarcely less incompatible; for it can have no application to that case at all, since an oscillator with one degree of freedom has no moment of momentum. The considerations of the next few paragraphs are very striking evidence for one of these two principles, and for the fundamental idea of Bohr, which underlies them both, namely that the possible states of the atom are determined by h and by nothing else. But they are not sufficient to decide between them.

The first consequence is that we find a meaning for Rydberg's constant N and can relate it to the other universal constants e, μ, h.

For, the assumption that (24) reduces to (26) is equivalent to the assumption that

$$N = \frac{2\pi^2 e^2 E^2 \mu}{h^3} \quad \text{.....(27).}$$

This result was, when the theory was first propounded, its most arresting feature. For many years it had been known that N was, at least very nearly, a universal constant, but nobody had any idea what it meant. It was now expressed simply in terms of other recognised universal constants, and the value calculated from these constants, $1{\cdot}091 \times 10^5 . c$, agreed well within experimental error with the observed value $1{\cdot}097 \times 10^5 . c$. But it should be noted that N has not really the form of a universal constant, because it contains E, the charge on the nucleus, which differs for different elements. However E is always an integral multiple of e, and N for any other element will be an integral multiple of N for hydrogen. In what follows, when we have reason to make a distinction, the element concerned will be denoted by a suffix and the N for hydrogen will be called N_H. When no suffix is attached N_H will be meant.

It will be convenient now to rewrite (17), (19), (20), introducing the value of N from (27). We find

$$W = W_0 - \frac{2\pi^2 \mu e^2 E^2}{h^2} \cdot \frac{1}{m^2} \quad \text{.....(17.1),}$$

$$a = m^2 \cdot \frac{h^2}{4\pi^2 \mu e E} \quad \text{.....(19.1),}$$

$$\omega = \frac{1}{m^3} \cdot \frac{8\pi^3 \mu e^2 E^2}{h^3} \quad \text{.....(20.1)}$$

where W_0 is the value of W when $m = 1$.

(19.1) gives a reasonable value for the radius of the atom in its normal state, which is that for which $m = 1$. For we have, putting in numerical values, $a = 0{\cdot}53 \times 10^{-8}$. The estimate of the radius of the hydrogen molecule (which is almost certainly larger than that of the atom) based on the study of gases is $1{\cdot}1 \times 10^{-8}$. The corresponding frequency of revolution $\frac{\omega}{2\pi}$ is $4{\cdot}13 \times 10^{16}$ sec.$^{-1}$, which is greater than that of any line of the hydrogen spectrum; but of course the theory does not indicate any close relation between frequency of revolution and frequency of radiation emitted.

The mass of the nucleus. To examine the theory further other spectra must be considered. There is only one other spectrum, known experimentally, which also arises from an atom containing one electron and a nucleus, and should therefore be strictly subject to the theory so far developed. This is the spark spectrum of helium. For the helium atom ($Z = 2$) in its normal state has two electrons, and, if the spark spectrum is that of an atom which has already lost one electron (see p. 23), it should have only one when emitting that spectrum. It will differ from the hydrogen atom in the charge E on its nucleus, which will be $-2e$ in place of $-e$; we must therefore substitute N_{He} for N_H. Since by (27) N is proportional to E^2, $N_{He} = 4N_H$.

If this were all, the lines of the helium spark spectrum would coincide with some of the hydrogen lines; for we should have for the terms, in place of (18),

$$\nu = N_{He}\left(1 - \frac{1}{m^2}\right) = N_H\left(4 - \frac{1}{\left(\frac{m}{2}\right)^2}\right) \quad \text{.....(28).}$$

The difference in the constant, 4 instead of 1, makes no difference in calculating the lines, and $m/2$ is an integer whenever m is even. Accordingly there should be a line of the helium spark spectrum coincident with every line of the hydrogen spectrum which is such that m is even for both of the terms involved in it. Since the constant term of the Balmer series has $m = 2$, alternate lines of this series should coincide with helium lines. There is actually in the spectra of certain stars a set of lines, known as the Pickering series, very nearly but not quite coincident with alternate members of the Balmer series, the remaining lines of that series being absent. Until recently these lines were attributed to hydrogen; Bohr attributed them to the helium atom which has lost one electron and explained why they are not exactly coincident with the hydrogen lines.

In establishing (27) we assumed that the mass of the nucleus was infinite compared with that of the electron. If the mass of the nucleus is finite and equal to M, then (14) and (15) give the path of an electron round a centre which coincides, not with the nucleus, but with the centre of mass of the nucleus and the electron.

If it is the nucleus that represents the observer's fixed point to which the motion should be referred, a well-known proposition of dynamics states that we must substitute for μ the quantity $\frac{M\mu}{M+\mu}$. Accordingly, in our numerical estimates we ought to take for μ, not the true mass of the electron, μ, but its apparent mass, $\mu' = \frac{M\mu}{M+\mu}$. Now M, the mass of the nucleus, is different for hydrogen and helium; for the latter it is about four times as great as for the former; and the ratio of N for the hydrogen atom to N for the helium atom ought to be, by (27),

$$\frac{N_H}{N_{He}} = \frac{E^2{}_H \mu'_H}{E^2{}_{He} \mu'_{He}} = \frac{1}{4} \cdot \frac{M_H (M_{He}+\mu)}{M_{He}(M_H+\mu)} \quad \text{.....(29).}$$

N_H and N_{He} can be measured with great accuracy. Paschen finds

$$N_H = 109677{\cdot}691^1,$$
$$N_{He} = 4 \times 109722{\cdot}144.$$

We know M_{He}/M_H from atomic weight determinations to be 3·99. Hence we find

$$\frac{M_H}{\mu} = \frac{4N_H - \frac{M_H}{M_{He}} \cdot N_{He}}{N_{He} - 4N_H} = 1850 \quad \text{.....(30).}$$

On the other hand, this ratio of the masses of a hydrogen nucleus and an electron can be calculated from

$$\frac{e}{\mu} = 5{\cdot}31 \times 10^{17} \quad \text{and} \quad \frac{e}{M_H} = 9570 \, . \, c$$

(given by the electro-chemical equivalent of hydrogen); the result is the value

$$\frac{M_H}{\mu} = \frac{e}{\mu} \Big/ \frac{e}{M_H} = 1850 \quad \text{.....(31).}$$

The agreement is perfect[2].

[1] The reason for the difference between this figure and that given previously, 109706, will appear on p. 86.

[2] It is rather too perfect and is partly due to the mutual cancellation of small experimental errors. For atomic weights give, not the ratio of the nuclei, but the ratio of the nuclei with their electrons—if it is true that the weights of nuclei and electrons are additive. It is unnecessary to enter into these refinements, but it should be noted that this method is probably the most accurate at the present

This numerical agreement of the values of M_H/μ is undoubtedly the most striking evidence in favour of Bohr's theory in its original form. There is no doubt left that the assumptions which are necessary to the calculation which leads to so remarkable a result must be embodied in any adequate theory of the origin of spectra. However it is not yet certain that the assumptions that have actually been made are necessary; they are sufficient—so much has been proved—but it is yet possible that some part of them may be altered or abandoned without changing the conclusions drawn from them.

X-ray spectra. The only other atomic systems which, according to Rutherford's theory, contain only one electron are the lithium atom which has lost two electrons, the beryllium atom which has lost three, and so on. But nothing is known of the spectra of these systems, although it is suspected that the spectrum of doubly-charged lithium may be present in some stars. Further evidence for Bohr's theory has, therefore, to be sought in systems possessing more than one electron, but in applying the theory to them new difficulties enter. For it is well known that our present mathematical analysis is incapable of predicting generally from the principles of dynamics the orbits of more than two bodies moving under their mutual forces. It can be shown that certain simple orbits are among those that are mechanically possible, but it can always be shown that there are also other orbits which are also mechanically possible and yet cannot be described analytically. The fundamental assumption that the possible states of the atom are to be found among the mechanically possible orbits cannot lead to any certain conclusions; for we can never know all the members of the class to which should be applied any further principle that may be

time for determining M_H/μ and e/μ. The value for the latter turns out to be $1{\cdot}7686 \times 10^7$. *c.*

It is worthy of note that the greatest source of experimental error is not the spectroscopic determination of N_H and N_{He}, although the whole difference $N_{He} - 4N_H$ is only 4 parts in 10,000. N can be measured to 4 parts in 10,000,000—such is the extraordinary sensitiveness of modern interference methods and such the accuracy of the theoretical equations. The greatest error is in the ratio M_{He}/M_H.

Most books of reference give $e/M_H = 9647$; but this value assumes that the atomic weight of hydrogen is 1 relative to Ag = 107·88; it is really 1·008.

suggested for selecting the possible states from the mechanically possible orbits. On the other hand it is worth while to examine those mechanical orbits which can be calculated, and to inquire whether possible states of the atom can be found among them which will satisfy the conditions imposed by the known terms of the spectrum. If we succeed in finding such states, the theory will be confirmed and light will be thrown on the structure of the atom; if we fail, the theory is not invalidated, because the possible states may yet lie among the mechanical orbits that cannot be calculated.

Proceeding in this manner a notable advance can be made in the interpretation of X-ray spectra. The excitation of X-ray spectra is always accompanied by the ejection of an electron from the atom or its displacement within the atom. The process is completely analogous to the excitation of light by ionisation, complete or partial, but the electron ejected or displaced is not one of the valency electrons on the surface of the atom[1], but one of the deep seated electrons near the nucleus. In the atoms of higher atomic number of which alone the X-ray spectra are known, the charge on the nucleus is many times that on an electron; moreover the field of force due to the electrons remote from the nucleus, if they are distributed symmetrically, will be very small, because the force due to one of these remote electrons will be nearly compensated by that due to the others. Accordingly, as a first approximation we may suppose that the force of an electron near the nucleus will be simply that due to the nucleus. On this assumption the mechanical orbits of such electrons are those which have already been discussed, namely circles or ellipses; once again we shall confine our attention to circles. The possible states of the atom, so far as these electrons are concerned, are then circles round the nucleus of which the radius, frequency of revolution, and energy are given by (17·1), (19·1), (20·1). E is now $-Ze$, where Z is the atomic number, and this value of E must be used in those equations.

Further, we concluded on p. 29 that the possible states concerned in X-ray spectra all represented positions (or motions) of electrons which were occupied in the normal state of the atom. We suppose therefore that there is an electron revolving in each

[1] 'Surface' may be too definite. It would be safer to say 'parts of the atom where the electric field due to the remaining charges is comparatively small.'

of the circles of which the radius is given by (19·1)—or at least in such of them as are so near the nucleus that they fall within the range of our assumption. And since orbits in which several electrons follow each other at equal intervals round the same circular orbit are mechanically possible, we must take into account the possibility that some or all of these circular orbits contain more than one electron. Proceeding to a second approximation we should take into account the forces between the various electrons in these orbits. It is easy to show that the effect on an exterior orbit (i.e. one with a larger m and larger a) of q electrons in the interior orbits is the same as that of diminishing the charge on the nucleus by qe; while the effect of increasing the number of electrons in any one orbit from 1 to p, is the same as that of diminishing the charge on the nucleus by s_p, where

$$s_p = \frac{1}{4} \sum_{k=1}^{k=p-1} \operatorname{cosec} \frac{\pi k}{p} \quad \text{.....(32).}$$

Accordingly, to this degree of approximation, the energy of the mth possible orbit should be given by writing in (17·1)

$$E = -e(Z - q - s_p);$$

so that

$$W = W_0 - \frac{hN_H(Z - q - s_p)^2}{m^2} \quad \text{.....(33).}$$

It follows that the variable quantity, depending on m, in the term of the spectrum is not $\frac{N_H Z^2}{m^2}$, but $\frac{N_H(Z - q - s_p)^2}{m^2}$, and that the line representing the transition from m_1 to m_2 should have the frequency

$$\nu = N_H\left(\frac{(Z - q_1 - s_{p_1})^2}{m_1^2} - \frac{(Z - q_2 - s_{p_2})^2}{m_2^2}\right) \quad \text{.....(34).}$$

Since q and s_p are both small compared with Z, (34) will be nearly equivalent to

$$\nu = N_H(Z - z)^2\left(\frac{1}{m_1^2} - \frac{1}{m_2^2}\right) \quad \text{.....(35),}$$

where z is some number intermediate between $q_1 + s_{p_1}$ and $q_2 + s_{p_2}$; z will depend on m_1 and m_2, but not on Z; it will be the same for all elements but will be different for the different lines in the spectrum. So far as dependence on Z is concerned, (35) agrees with the relation (11) found experimentally.

By the careful examination of z it should be possible, if the theory is correct, to determine what are the values of m corresponding to the orbits immediately surrounding the nucleus, and what is the number of electrons in each ring; for z is a function of these numbers. But an exact determination is difficult, partly because the theory given, professedly only an approximation, is not quite true even with respect to the considerations so far introduced, partly because there are other complications arising from matters considered on p. 89. But it seems certain that the innermost orbit corresponds to $m = 1$, and that there is only one electron in this ring. The lines of the K-spectrum are produced by transitions to this orbit from outer orbits; and the orbit is therefore known as the K-ring. The next orbit has $m = 2$ and there are several electrons in it—perhaps eight, but this figure is somewhat speculative; the lines of the L-spectrum are produced by transitions to this ring from outer rings, and it is known as the L-ring. The next ring is doubtless the M-ring, connected in the same manner with the M-spectrum; it almost certainly has $m = 3$, but still less is known about the electrons in it. The fact, indicated in Fig. 2, that z increases from K- to L-spectrum and from L- to M-spectrum is, of course, in accordance with this view that the number of electrons in each of the possible orbits increases as we proceed in this direction.

But now we should be able to use (35) to calculate the frequencies of the lines. If the K-ring has $m = 1$ and the L-ring $m = 2$, the value of $\frac{1}{m_1^2} - \frac{1}{m_2^2}$ for the line corresponding to a transition between these rings should be $\frac{3}{4}$, and A in (11) for this line should be $\frac{3}{4}$. The square of the slope of the straight line representing K_α in Fig. 2 actually gives a value 0·76 for A, and this line may therefore be identified confidently with this transition. Again there should be a line in the L-spectrum with $A = \frac{1}{2^2} - \frac{1}{3^2} = \cdot 139$; A for L_α found from Fig. 2 is ·138. Again there should be a line in the M-spectrum with $A = \frac{1}{3^2} - \frac{1}{4^2} = \cdot 0486$; from Fig. 2, A for M_α is ·0445 but here doubtless the degree of approximation that has been

adopted is becoming insufficient, because the distance from the nucleus and the effect of the mutual actions of the electrons are greater. These facts were pointed out by Moseley, who interpreted the relations that he had discovered in accordance with Bohr's theory, and thus founded the science of X-ray spectroscopy. The concordance between the simple theory and the experimental facts is sufficient to show beyond doubt that the general ideas underlying the theory are correct, and that optical and X-ray spectra are manifestations of precisely the same principles. In fact, we may almost say that the optical spectrum of hydrogen is its X-ray spectrum, that the terms corresponding to $m = 1, 2, 3$ in Fig. 1 are the K-, L-, M-terms, and that the Lyman series is the K-spectrum, the Balmer the L-spectrum, the Paschen the M-spectrum. Almost, but not quite! For in the atoms of higher atomic number, which give typical X-ray spectra, the L- and M-rings are normally occupied by electrons; in the hydrogen atom only the K-ring is occupied, for there is only one electron. This observation suggests a further test. If we trace the L- or M-spectrum down to very low values of Z, the points should suddenly break away from the lines of Fig. 2 when Z, the number of electrons in the atom, becomes too small for the L- or M-ring to be occupied. Moreover when the M-ring has completely disappeared, the electrons in the L-ring become the valency electrons on the 'surface' of the atom; the L-absorption limit should be the term $(1, s)$ of the optical spectrum and the corresponding energy be that required to ionise the atom. Kossel has followed up these suggestions[1], and has found that they do accord with experiment; but the aspect of the matter belongs more properly to the study of the constitution of the atom, to which the study of spectra in the light of Bohr's theory is one of the most certain clues.

Other optical spectra. It is unfortunately impossible to attain the same success in the interpretation of the remaining optical spectra. In the original exposition of his theory Bohr suggested that all the electrons in the atom might be revolving in rings, all in the same plane, similar to those on which the innermost electrons

[1] Since in the critical range the spectra usually lie outside the region accessible to direct experiment, Kossel has used (13) and similar relations to find the frequency of very soft L- and M-rays.

appear to lie. But this suggestion has had to be abandoned; it is almost certain now that the outermost electrons, and especially the valency electrons, do not lie in one plane. This conclusion is in no way inconsistent with the essential basis of Bohr's theory, for coplanar rings only represent one class of mechanical orbits; but even if the other orbits could be calculated, the theory, in the form in which it has been expressed so far, provides no principles by which the possible states may be selected from the mechanical orbits.

Nevertheless there is one feature common to all optical spectra to which the simplest form of the theory gives an immediate interpretation, namely the occurrence in the terms of all of them of the constant N_H. For if we compare (8) or (9), representing the terms of any series spectrum, with (7) representing those of hydrogen, we see that they tend to identity as m becomes very large. On the other hand, according to (19) a increases as m^2, and with increasing m the radius of the orbit rapidly becomes large compared with the value for $m = 1$, which represents roughly the radius of the atom in its normal state; for the radii of all atoms are not very different from that of hydrogen. Now when an electron separated from an atom is removed to a distance from it large compared with its radius, the forces on it will be very nearly those which would be exerted by a particle, carrying the same total charge as the atom, situated at its centre. But the charge on any atom from which one electron has been separated is $-e$, which is the value of E for the nucleus of the hydrogen atom. Accordingly the orbits of the electron, which are at a considerable distance from the atom, and the possible states of the atom corresponding to various conditions of this electron should be very nearly the same as those of an electron revolving round a hydrogen nucleus. We thus find a simple and immediate explanation of the fact that, when m is large, the terms of the spectra of all atoms approach closely to those of the hydrogen spectrum. The fact that N in Rydberg's formula is a universal constant means that all systems with the same total charge exert the same force on electrons very distant from them.

This statement should apply only to the flame spectra which represent states intermediate between the normal and that in which

one electron is lost. The spark spectra, representing states up to that in which two electrons are lost, should approach in the same manner the spark spectrum of helium. Such is actually found to be the case; in spark spectra, N or N_H is replaced by $4N$ or N_{He}; that fact, which had been definitely established before Bohr's theory was propounded, provides a new confirmation for the theory[1].

Bohr's Theory and mechanics. Such is the evidence for the assumptions characteristic of the second part of Bohr's theory. Is there any evidence to set against it? So far as I know, only one experimental fact has been pointed out that seems, even at first sight, to conflict with it. According to (19) the radius of the orbit corresponding to the term $m = 30$ should be nearly 1000 times the normal radius of the atom, i.e. about 10^{-5} cm. Lines involving this term have been observed when the distance between the molecules was no greater than this radius; and it has been urged that under such circumstances the orbit could not be maintained in view of collisions with other atoms. Nobody will be disposed to put too much weight on such an argument—though it should not be overlooked—and more serious objections can be based on the intrinsic limitations or even inconsistencies of the theory.

One difficulty may seem to the reader to have been treated too lightly on p. 43. It may appear inconsistent that a theory which sets out to replace mechanics should start by making use of mechanical principles. This apparent inconsistency can be resolved; for mechanical principles are known to be true only within a certain range, and within that range Bohr's principles agree with them. Thus mechanics asserts that the frequency of radiation is that of the vibrations of the system from which it arises; Bohr's theory denies that proposition for atomic systems, but accepts it for the quite different systems to which mechanics is known to apply. From (18) it follows that the frequency of the radiation emitted

[1] It will be noted that the correction discussed on p. 49, arising from the finitude of the mass of the nucleus, is left out of account. There is no hope of establishing the need for this correction, for, as was noted before, neither (8) nor (9) represents with perfect accuracy the terms of the spectrum; there is, therefore, a limit to the accuracy with which N can be deduced from the spectrum.

in a transition from an orbit characterised by m_1 to an orbit characterised by m_2 is given by

$$\nu = \frac{N}{m_1^2} - \frac{N}{m_2^2} \quad \text{.....(36).}$$

If m_1, m_2 are very large compared with the difference between them, we may put $m_1 = m_2 = m$ in the denominator, and (36) reduces to

$$\nu = 2N \cdot \frac{m_1 - m_2}{m^3} \quad \text{.....(37).}$$

On the other hand the frequency of revolution, given by (20) after substituting for N from (27), is

$$\frac{\omega}{2\pi} = 2N \cdot \frac{1}{m^3} \quad \text{.....(38).}$$

That is to say, according to Bohr's theory, when m is very large the frequency of the radiation emitted in a transition for which $m_1 - m_2 = 1$ is equal to the frequency of revolution of the electron and therefore equal to that predicted by mechanical principles. The only assumptions that are needed to remove all contradiction between the older and the newer theory are (1) that all radiation emitted under conditions in which the older theory is known to be applicable is emitted in transitions between states for which m is very large, (2) that in these transitions m only changes by one unit. The second assumption will be discussed later; the first is obviously true. For if we know the charges on attracting particles, their masses and the radius of their orbit, we can calculate m from (19) and (27). If we choose these quantities to be the smallest possible for any system the motion of which and the radiation from which can be investigated by direct 'mechanical' experiment, we shall find that the most sanguine estimates of sensitive experiment lead to values of m very much greater than a million.

The view that we should take of the apparent inconsistency which is under discussion is not that Bohr's theory introduces at one stage of the argument principles that it subsequently rejects, but that Bohr's theory as a whole is a statement of the general case, and that the 'mechanical' theory is applicable only to limiting conditions. The familiar mechanical calculations which determine the class within which possible states of a radiating system lie are

an integral part of the newer theory; but this theory asserts also that not all the states in that class are possible states, as was formerly believed. It is only in limiting conditions that all the states in the class are possible states and that the older theory becomes identical with the newer.

There remains, however, one much more serious objection to Bohr's theory in the form in which it was originally presented. It is only applicable to circular orbits performed under central forces. The form of the characteristic assumption (26) suggests that it might be applied equally to all orbits in which the moment of momentum is constant, that is to all orbits (whether circular or not) performed under any kind of central force. But if we attempted to apply it even to the case most nearly resembling that which we have considered, namely the elliptical orbits of an electron round a nucleus, we should not actually attain results confirmed by experiment—as we shall see presently. We might simply deny that any but circular orbits were possible, but that assumption is not intrinsically plausible; moreover, it would leave totally unexplained some features of spectra that have been left out of our recent discussion, for example, the fact that the hydrogen and helium lines are not perfectly simple, but are made up of several components and satellites. We must seek therefore to extend the theory and to replace (26) by some more general assumption, which will reduce to (26) when the orbits are circular, but will give results in accordance with experiment when they are not. Such an assumption can be found. And from the fact that it can be found, and that the simple assumption from which we started must be replaced by it, we must not conclude that Bohr's original theory is worthless. It merely bears the same relation to the more general theory that Newtonian bears to Hamiltonian mechanics; it is the necessary first stage of the successive generalisations by which alone the ultimate and complete theory can be attained.

V. LINES AND COMPONENTS

The generalisation of Bohr's assumption. The extension of Bohr's principle to all orbits that are periodic in the most general sense is due to Sommerfeld.

In the form in which dynamical principles were stated by Hamilton (following Lagrange), the state of the system under consideration at any instant is supposed to be described by n coordinates $q_1, q_2, \ldots, q_k, \ldots, q_n$. The q's need not be coordinates in the sense understood in analytical geometry (though of course such spatial coordinates are included in the meaning of the term); but they must be such that, when the values of the q's at any instant are given, the state of the system at that instant is completely fixed. If n is chosen to be as small as possible, consistently with this condition being fulfilled, it is called the number of degrees of freedom of the system. The motion, or change of state, of the system at the same instant is then completely described by the n quantities $\dot{q}_1, \ldots, \dot{q}_n$. The energy of the system W (apart from a physically insignificant constant term) can be expressed as the sum of two terms: $W_{\text{pot.}}$, the potential energy, which is a function of the q's and not of the $\dot{q}$'s, and $W_{\text{kin.}}$, the kinetic energy, which is a function of both q's and $\dot{q}$'s. In all the cases with which we shall be concerned, $W_{\text{kin.}}$ is a homogeneous quadratic function of the $\dot{q}$'s, being of the form $\Sigma a \dot{q}_r \dot{q}_s$, where r and s may be the same, and the a's are either constants independent of the time or are functions of the q's which vary with the time.

In place of the $\dot{q}$'s, we may use to describe the state of motion quantities known as 'momenta' or 'impulse coordinates,' $p_1, \ldots, p_n$. The p's are related to the q's and $\dot{q}$'s by the equation

$$\dot{p}_k = \frac{\partial W_{\text{kin.}}}{\partial \dot{q}_k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (40).$$

Then, according to Hamilton, the motion of the system, starting

from any prescribed conditions, will be determined completely by the equations

$$\frac{dq_k}{dt} = \frac{\partial W}{\partial p_k} \quad \text{.....(41);}$$

$$\frac{dp_k}{dt} = -\frac{\partial W}{\partial q_k} \quad \text{.....(42),}$$

where in the partial differential coefficients every variable except that written below the line is constant.

A very simple example may make the matter clear to those unfamiliar with analytical dynamics. The state of a falling particle at any instant can be completely described by x, its distance below the starting point, and its motion by $\dot{x}$. The energy, expressed in the Hamiltonian form, is

$$W = W_{\text{pot.}} + W_{\text{kin.}} = -mgx + \tfrac{1}{2}m\dot{x}^2 \quad \text{.....(43).}$$

Consequently

$$p = \frac{\partial W_{\text{kin.}}}{\partial \dot{x}} = m\dot{x} \quad \text{.....(44).}$$

(41) and (42) then become

$$\dot{x} = \dot{x} \quad \text{.....(45);}$$

$$m\ddot{x} = mg \quad \text{.....(46).}$$

The first equation reduces to an identity because the most general choice of coordinates consistent with the conditions has not been made. But (46) states in the Newtonian form the complete specification of the motion.

This is not the place to inquire into the evidence for the Hamiltonian equations. It is sufficient to say that it can be shown generally that, whenever the system consists of particles subject to Newtonian forces, the Hamiltonian form of the principles of dynamics reduces to the Newtonian. The Hamiltonian form is valuable, partly because it enables us to deal with systems that cannot be analysed into Newtonian particles, partly because it is mathematically convenient. But neither of these advantages concerns us immediately; for all the systems that will be considered can be analysed into Newtonian particles, and we shall avoid all analytical difficulties. On the other hand, since Sommerfeld's principle cannot be stated apart from the method of describing the

system adopted by Hamilton, it is well to understand the primary purpose of that method.

Sommerfeld's generalisation of the original assumption made by Bohr may now be stated thus. If the motion of the conservative system is periodic, the only states of the system which are possible, in the sense of the previous discussion, are those for which

$$\int p_1 dq_1 = m_1 h\,; \quad \ldots\ldots\,; \quad \int p_k dq_k = m_k h\,; \quad \ldots\ldots \quad \ldots(47).$$

In these equations, of which one applies to each of the n coordinates with its associated momentum, the integral on the left-hand side is to be taken over the complete range through which p_k and q_k pass in a period; on the right-hand side, h is Planck's constant, and $m_1, \ldots, m_k, \ldots$ are integers which may have different values according to the different suffixes. Unless the motion is one for which n such conditions are fulfilled for the n coordinates, then, although it may be one consistent with the energy of the system, it will not be one that is possible as a state from or to which a radiating system can pass with the emission or absorption of homogeneous radiation.

Sommerfeld's assumption is stated here without some of the modifications and additions which will have to be mentioned later; but for the present it will be well to consider it in this form. It must be observed at once that there is an ambiguity in it. For the coordinates used to describe any system can always be chosen in several ways; for example, we may use rectangular or polar coordinates to fix the position of a particle. For the same motion the quantities $\int p dq$, taken over the same period, will differ according to the choice of coordinates; if rectangular coordinates are chosen, these integrals will have one set of values for a given motion, if polar coordinates are chosen another set[1]. Accordingly the motions which will satisfy the conditions (47), and will be selected as possible states, will differ with the choice of coordinates; the principle will give a definite result only if there is added to it some method of determining which coordinates are to be used. A perfectly general way of determining this matter has not yet been

[1] On the other hand, it can be shown that the sum of all these integrals is independent of the choice of coordinates; and it was doubtless this feature which helped to suggest his principle to Sommerfeld. But it is not relevant to our discussion.

suggested, but a way that applies to all the cases that we have to consider has been found. It will be described on p. 72; for the present the rule may be stated roughly by saying that the 'natural' coordinates are to be chosen, that is to say the coordinates which anyone would naturally choose for describing the system on the grounds of analytical simplicity. Thus, nobody would think of choosing for the description of motion in a straight line any coordinate but distance from a point in that line; and nobody would think of describing motion in a circle by any coordinate but angular distance from some radius; any other choice would be artificial. This rule must suffice for the present. But anyone inclined to think it so vague as to deprive the principle (47) of any significance, should observe that, even if no rule could be given at all, and even if the choice had to be sought for by trial and error in every application, it would still be a significant and important fact that it is always possible to choose some coordinates such that, in conjunction with (47), they actually lead to results confirmed by experiment. For that fact could never have been anticipated by one who had not heard of Bohr's theory and would never have been discovered if the fundamental ideas on which it is based had not been propounded.

Two simple examples. Our first justification of the principle must be to show that it reduces to (26) when applied to the circular orbits considered by Bohr. Here, as was said, the natural coordinate is ϕ, the angular position of the particle, which is such that $\dot{\phi} = \omega = \text{constant}$. Then

$$\left.\begin{aligned} W_{\text{kin.}} &= \tfrac{1}{2}\mu a^2\dot{\phi}^2 \\ p_\phi &= \frac{\partial W_{\text{kin.}}}{\partial \dot{\phi}} = \mu a^2 \dot{\phi} \\ \int_0^{2\pi} p_\phi d\phi &= 2\pi\mu a^2\omega \end{aligned}\right\} \quad \text{.....(48),}$$

and (47) reduce to

$$\mu a^2 \omega = m \cdot \frac{h}{2\pi},$$

which is (26).

Further it was noted on p. 47, as a slight objection to Bohr's principle (26), that it did not agree with that which is the basis

of Planck's theory of the linear oscillator. We now find that Planck's principle is, like Bohr's, a consequence of Sommerfeld's more general principle, which includes both. For, in a linear oscillator, we take x as our coordinate; p_x is then $\mu\dot{x}$, where μ is again the mass of the particle. But $x = A\cos\frac{2\pi t}{T}$, where $T = \frac{1}{\nu}$ is the period. (47) then reduce to

$$\mu\int_0^T \dot{x}\,dx = \frac{2\pi^2 A^2\mu}{T} = mh \quad \text{.....(49).}$$

But W, the energy of the oscillator, reckoned at the moment when it is also kinetic is given by

$$W = \tfrac{1}{2}\mu\cdot\left(\frac{2\pi A}{T}\right)^2 = \frac{2\pi^2 A^2\mu}{T^2} \quad \text{.....(50),}$$

so that

$$W = \frac{mh}{T} = mh\nu \quad \text{.....(51),}$$

which again agrees with Planck's assumption.

Accordingly Sommerfeld's assumption includes as special cases both that of Planck and that of Bohr; and this reduction of two apparently contradictory assumptions to a single more general principle is by itself no mean achievement. But the great advantage of the new principle is that it is far more widely applicable.

Elliptical orbits. For we can now investigate elliptical orbits as well as circular; and it must clearly be our next task to inquire whether the recognition of such orbits is consistent with the theory of spectra.

We again take polar coordinates, and now naturally choose the focus as origin. Then $q_1 = r$, $q_2 = \phi$; and we find easily from (44) $p_r = \mu\dot{r}$, $p_\phi = \mu r^2\dot{\phi}$. (47) become

$$\int\mu\dot{r}\,dr = m_1 h \quad \text{.....(52);}$$

$$\int\mu r^2\dot{\phi}\,d\phi = m_2 h \quad \text{.....(53),}$$

the integrals being taken round the ellipse. $\mu r^2\dot{\phi}$ is constant throughout the elliptical orbit as throughout the circular; let its constant value again be p; then from (53)

$$\mu r^2\dot{\phi} = p = m_2\cdot\frac{h}{2\pi} \quad \text{.....(54).}$$

(54) is, of course, identical with (26), but we have now to consider also (52). Here the calculation is rather more complex, and the analysis will only be sketched. The equation of the ellipse is

$$\frac{a}{r}(1-\epsilon^2)=1+\epsilon\cos\phi \quad \ldots\ldots\ldots\ldots\ldots\ldots(55),$$

where a is the major axis and ϵ the eccentricity. Hence

$$\frac{1}{r}\frac{dr}{d\phi}=\frac{\epsilon\sin\phi}{1+\epsilon\cos\phi} \quad \ldots\ldots\ldots\ldots\ldots\ldots(56).$$

But

$$\dot{r}\,dr=\dot{\phi}\left(\frac{dr}{d\phi}\right)^2 d\phi,$$

so that from (54)

$$\int\mu\dot{r}\,dr=\int_0^{2\pi}p\epsilon^2\frac{\sin^2\phi\,d\phi}{(1+\epsilon\cos\phi)^2}=2\pi p\left(\frac{1}{\sqrt{1-\epsilon^2}}-1\right) \quad \ldots(57).$$

Accordingly (52) becomes

$$2\pi p\left(\frac{1}{\sqrt{1-\epsilon^2}}-1\right)=m_1 h,$$

or, from (54),

$$1-\epsilon^2=\frac{m_2^2}{(m_1+m_2)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(58).$$

(58) imposes a general limitation on the elliptical orbits that are possible; values of ϵ can occur only if they are such that $1-\epsilon^2$ is the ratio of two square integers. (58) in conjunction with (54) imposes further a limit on the dimensions of the orbits that can occur having these permissible eccentricities. The permissible values of the major axis, a, can be calculated in the following manner, which gives us at the same time the values of the energy for the permissible orbits:

$$W_{\text{kin.}}=\frac{\mu}{2}(\dot{r}^2+r^2\dot{\phi}^2)=\frac{p^2}{2\mu r^2}\left\{\left(\frac{1}{r}\frac{dr}{d\phi}\right)^2+1\right\}$$

$$=\frac{p^2}{\mu a^2(1-\epsilon^2)^2}\left(\frac{1+\epsilon^2}{2}+\epsilon\cos\phi\right) \quad \ldots\ldots\ldots\ldots(59),$$

$$W_{\text{pot.}}=-\frac{eE}{r}=-\frac{eE}{a}\cdot\frac{1+\epsilon\cos\phi}{1-\epsilon^2} \quad \ldots\ldots\ldots\ldots(60).$$

Now $W_{\text{kin.}}+W_{\text{pot.}}=W$ must be independent of ϕ, since it is

constant round the orbit; hence the terms in (59) and (60) involving $\cos\phi$ must cancel, which gives

$$a = \frac{p^2}{\mu e E\,(1-\epsilon^2)} = \frac{h^2}{4\pi^2 \mu e E}(m_1+m_2)^2 \quad \ldots\ldots\ldots(61)$$

and

$$W = W_{\text{kin.}} + W_{\text{pot.}} = \frac{eE}{a(1-\epsilon^2)}\left(\frac{1+\epsilon^2}{2} - 1\right) \quad \ldots\ldots(62).$$

If the values of a and of ϵ are taken from (61) and (58), and if the energy is reckoned from the normal state as in (17·1), we have

$$W = W_0 - \frac{2\pi^2 \mu e^2 E^2}{h^2} \cdot \frac{1}{(m_1+m_2)^2} \quad \ldots\ldots\ldots(63).$$

The possible elliptical orbits are now described completely, as were the possible circular orbits by (17·1), (19·1), (20·1).

Total and partial quantum numbers. Our conclusions may be stated in words thus. The collection of possible states of the atom with one electron and a nucleus is characterised by two integers or 'quantum numbers' m_1 and m_2; in virtue of their connections with r and ϕ, they are usually termed respectively the 'radial' and 'azimuthal' quantum numbers; either of them will be termed a 'partial' quantum number, and $m = m_1 + m_2$ will be termed the 'total' quantum number. The collection includes both circles ($\epsilon = 0$) and ellipses ($0 < \epsilon < 1$), and forms a double series. First, there is a series of states with different total quantum numbers. According to (61) and (63) these states have different major axes and different energies, the axis and the energy increasing with the total quantum number. Second, each of the states of this series is made up of a group of subordinate states, all of which have the same total quantum number, but different partial quantum numbers; they all have the same major axis and energy, but have different eccentricities. There are m such subordinate states with total quantum number m, namely those for which $m_1 = 0, 1, 2, \ldots (m-1)$, and $m_2 = m, (m-1), (m-2), \ldots, 1$. Of these one is a circular orbit ($m_1 = 0$), while the remainder are elliptical orbits. The orbit ($m_1 = m$, $m_2 = 0$) is excluded because when $m_2 = 0$ and $\epsilon = 1$, the ellipse degenerates into two coincident straight lines passing through the focus, where is the nucleus. An electron vibrating in such an

orbit would have to pass through the nucleus; and we may expect that such an orbit will be impossible. Accordingly we shall assume that the azimuthal quantum number can never be 0.

And now what of the lines which should be emitted in transitions between these states? If we compare (63) with (17·1) or (61) with (19·1), we see that they differ only in the substitution of $(n_1 + n_2)^2$ for n^2. But n_1, n_2, n are all integers and n_1 may have the value 0. Therefore to every possible value of $(n_1 + n_2)^2$ will correspond an equal value of n^2; each of the circular orbits just calculated will be identical with one of the circular orbits calculated before, as was to be expected; it follows from what has just been said that every possible energy of an elliptical orbit is a possible energy for one of the circular orbits calculated before. But the frequencies of the lines in the spectrum are determined by the energies of the possible orbits between which the electron passes in the emission or absorption of those lines. We conclude then that the lines emitted or absorbed in passing between two possible elliptical orbits (or between an elliptical and a circular orbit) are exactly the same as the lines emitted or absorbed in passing between two circular orbits. The recognition of the elliptical orbits makes no addition to the lines to be expected in the spectrum, because the frequency of the lines is determined wholly by the total quantum number and is independent of changes in the partial quantum numbers so long as the total number is the same.

This conclusion is necessary if the theory is to be true; for in the hydrogen spectrum and in that of the charged helium atom, to which alone the theory is strictly applicable, all the observed lines have been accounted for on the assumption of circular orbits. If elliptical orbits introduced any new lines, they would be lines that have not been observed, and the theory would be weakened and not strengthened. Nevertheless this may seem a disappointing result of our calculations, for if the introduction of elliptical orbits accounts for no new facts, what is the object of introducing the new principles and how are they to be tested? One answer to the first question is that, though the recognition of elliptical orbits provides no new evidence for the theory, it removes a difficulty; for, as was remarked before, it was difficult to give any satisfactory reason why elliptical orbits should not be possible. If elliptical

orbits had been treated by the simple and original theory of Bohr, the quantum condition (54) would have been introduced into (62), but not the further condition (58). It would have been concluded that

$$W = W_0 - \frac{2\pi^2 \mu e^2 E^2}{m_2^2 h^2}(1 - \epsilon^2) \quad \ldots\ldots\ldots\ldots\ldots\ldots (64);$$

and since all values of ϵ, forming a continuous series, would have been admissible, the possible states would have formed a continuous series of which the energies differed infinitesimally. The theory would not have predicted sharp lines, separated by finite intervals, but broad bands some of which would have merged into others. It is only the introduction of the second quantum condition, arising from Sommerfeld's principle but not from Bohr's, that enables the existence of elliptical orbits to be reconciled with the most characteristic feature of a spectrum.

Resolution into components. But there is another and much more important answer to both questions. The calculations that have just been given show that a single line in the spectrum does not correspond to the passage from one single possible state to another single possible state; it corresponds to a passage from one group of possible states to another group, the states within each group being the collection of elliptical orbits (including one circular) all of which have the same total quantum number. In the normal hydrogen and helium spark spectra, and with the dynamical basis that we have adopted, the energies of all the states in one group have exactly the same value; but a very slight modification either of the structure of the atom, or of the circumstances in which it is placed, or in the assumptions made in the calculation will make these energies not exactly equal; and if they cease to be exactly equal we shall replace each single line of the spectrum by a group of lines, corresponding to the passage between each of the slightly differing states in one group and each of the slightly differing states in the other. The lines of the spectrum will no longer be single, but will be split up into a set of very slightly separated components.

This possibility suggests an explanation of facts which have so far been outside the scope of our theory. For it is known that in certain circumstances the single lines of a spectrum are split up

into several slightly separated components. Thus the well-known Zeeman effect, the discovery of which was the starting point of the whole electronic theory of spectra, consists in the splitting of a line into such components by the action of a magnetic field. Again Stark and Lo Surdo have recently found that a somewhat similar splitting occurs in a strong electric field. And lastly, it has long been known that very few, if any, spectral lines are really perfectly simple even in the absence of an external field; the lines of the Balmer series and of the helium spectrum are found on close examination to have a 'fine structure' and to consist of several components that can be distinguished under very high dispersion. The considerations that have just been put forward inevitably suggest that these components may represent the different lines, associated with different possible states in the same group, which are coincident if all the conditions we have assumed are strictly fulfilled, but are separated as soon as the conditions are slightly changed. This suggestion proves to be true and leads to the most remarkable confirmations of the theory we are considering.

Conditionally periodic orbits. We shall therefore now proceed to inquire how the possible states of the atom will be changed by the presence of a magnetic or electric field and whether we can find any other kind of 'perturbation' of the orbits which will explain the fine structure of the lines even where such a field is absent. It will be well to start with a few observations applicable to all the forms of the perturbation we are to consider.

The perturbed orbits will not be periodic in quite so simple a manner as the unperturbed. The orbits that we have considered so far are closed paths such that the electron constantly returns to exactly the same point at equal intervals of time. But such orbits are very particular examples of the general class of periodic orbits characteristic of systems with more than one degree of freedom. The more general examples (for two degrees of freedom) are illustrated in Figs. 3, 4. Fig. 3 represents a 'Lissajou' figure, drawn by the familiar mechanism of a pen resting on a plane oscillating in two perpendicular directions; Fig. 4 represents the motion of a planet in a path which is approximately an ellipse but of which the perihelion position is changing.

All these more general orbits, which are known as 'conditionally periodic,' have common features important for our purpose. Each of the coordinates of the orbit oscillates constantly between definite limits, but the times at which the limits are attained are different for different coordinates. Thus in Fig. 3, if we take x and y as coordinates, both x and y are constantly oscillating between the limits 0 and a, where a is the side of the circumscribing square; but x and y do not usually attain their limits at the same time,

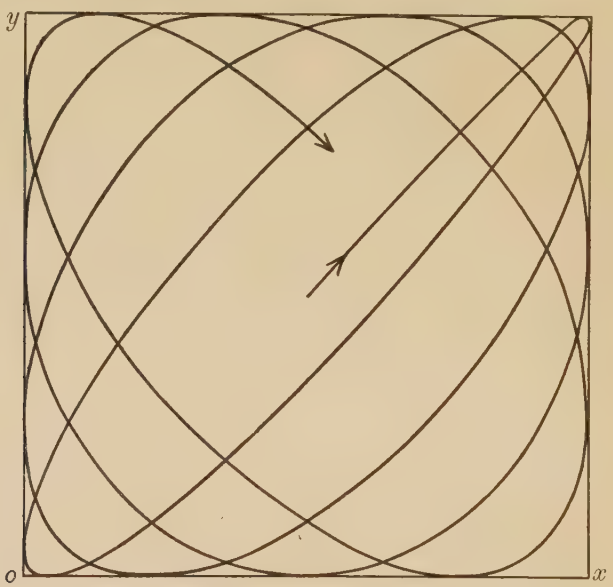

Fig. 3.

and the value of y at the instant when x attains its limit is constantly changing. Similarly in Fig. 4, if we take r and ϕ with origin at 0, r is constantly oscillating between the limits R_1 and R_2, the radii of the inscribed and circumscribed circles, while ϕ constantly oscillates between 0 and 2π; but at successive oscillations r does not attain its limits for the same values of ϕ. In both examples the orbits eventually cover the whole of a surface, the whole of the square in Fig. 3, and the whole of the surface between the circles in Fig. 4; in course of time the system will always return indefinitely nearly to its original position, but will not in

general regain within any finite time a position through which it has once passed. These features, suitably generalised for a greater number of coordinates, are characteristic of the conditionally periodic orbits of a system of any number of degrees of freedom. The purely periodic systems that we have discussed hitherto are

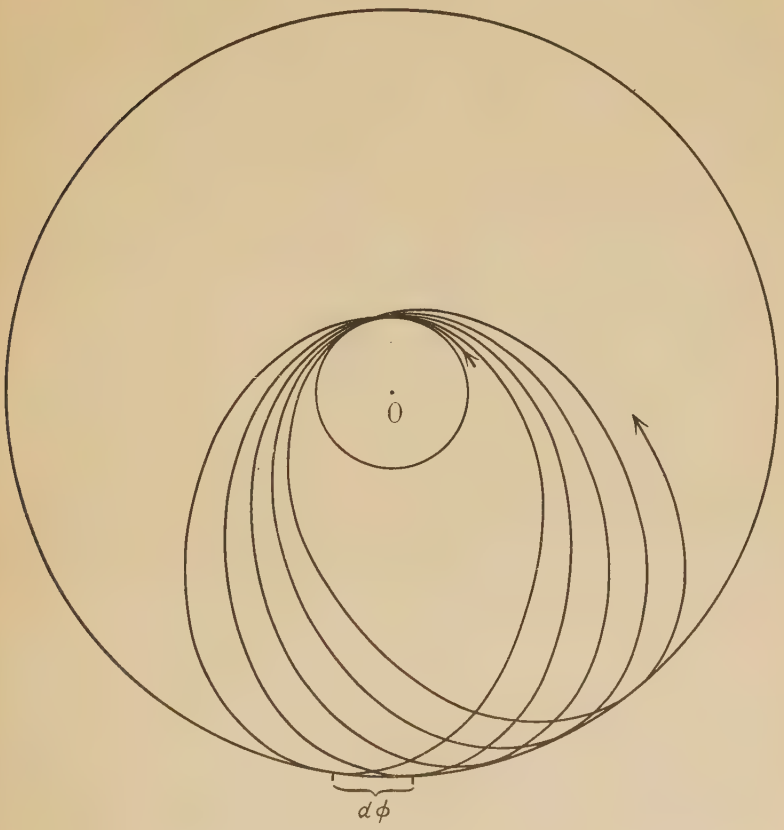

Fig. 4.

degraded examples of such orbits which only occur in very exceptional circumstances. Thus, the actual orbits of the planets round the sun are not strictly Keplerian ellipses; owing to mutual perturbations, they are really conditionally periodic orbits, differing but little from truly periodic orbits. Such orbits have been the

subject of very careful study because of their astronomical importance; they are equally important in the theory of spectra, for all the perturbed orbits that we are about to consider are of this type.

It will be seen readily that in these conditionally periodic orbits there is one specially distinguished set of coordinates that it is natural to take for describing the orbit. This set will be the coordinates of which the oscillation between fixed limits makes the orbit conditionally periodic. Thus, if in Fig. 3 we took polar coordinates with the point 0 as origin it would not be true that either r or ϕ oscillated between fixed limits; neither would start from one value, change continuously to another value, return to the first value through the same series of values as before, and repeat this process indefinitely; nor, if in Fig. 4 we took rectangular axes, would x and y oscillate in this manner. In Fig. 3 it is only if we take x, y as coordinates that we find this oscillation; in Fig. 4 it is only if we take r, ϕ. In these orbits there is an obvious method of removing the ambiguity as to coordinates mentioned on p. 63, and laying down definitely what set of coordinates is to be selected in the application of the fundamental principle formulated in (47). This method is actually adopted. It is part of Sommerfeld's principle that the coordinates to be chosen in dealing with conditionally periodic orbits are these special oscillating coordinates. And at the same time a difficulty that might arise about the limits of the integrals in (47) is resolved. So far it has been stated that the integrals are to be taken over a complete period; but in a conditionally periodic system there is no definite period; the system never returns to exactly the same state. The new rule, which follows obviously from the choice of coordinates, is that the limit of the integral $\int p dq$ is to be taken over a complete oscillation of the coordinate q. Thus in Fig. 3 the integral, when q is x, would be taken over a passage from $x = 0$ to $x = a$ and back again to $x = 0$. (It is easily proved that the integral over an oscillation is the same for all oscillations of the same coordinate, even though the other coordinates change in successive oscillations.)

This rule for the choice of coordinates is usually stated in a different and much less direct manner, involving analytical and not geometrical considerations. The special coordinates that have to be chosen are called 'separable,' because they are distinguished

analytically from other sets of coordinates by the fact that the expression for the momentum corresponding to each of these separable coordinates can be put in a form in which it contains only that coordinate together with constants of the system, the same for all coordinates; the expression in this form does not contain explicitly the other coordinates. (Cf. p. 100.) This property of the separable coordinates is of great importance in the analysis involved in deducing the motion from Hamilton's equations, but if it is described in this analytical form, its physical significance is concealed. Epstein has pointed out that, in all the cases with which the theory of spectra has been concerned so far, the analytical and the geometrical descriptions of the special coordinates agree. It should be observed that there is no proof that there must always be only one set of special coordinates having this property; indeed in some of the similar cases there would appear to be more than one set. But it seems that if there are two sets, they lead to the same possible states when the fundamental quantum relations (47) are applied to them. This proposition has not been proved generally, but no infraction of it has been discovered in the cases important for our purpose.

The Zeeman effect. And now let us proceed to turn to the three types of perturbation that have been indicated, and begin with the Zeeman effect.

The Zeeman effect is the one feature of spectral emission of which the older theory was able to offer a satisfactory explanation. (See Chap. VI.) From a consideration of the effect of a magnetic field upon the motion of a charged body, it could be shown that the light emitted from an electron vibrating harmonically in the absence of a field with frequency ν_0 should be changed by the presence of a magnetic field of intensity H into three components of frequency $\nu_0 + d\nu$, ν_0, $\nu_0 - d\nu$, where

$$d\nu = H \cdot \frac{e}{4\pi\mu c} \quad \text{.....(66).}$$

If the light is observed along the direction of the field, only the components $\nu_0 \pm d\nu$ should be observed and they should be circularly polarised in opposite senses; if it is observed at right

angles to the field, all three components should be observed and should be plane polarised. This 'normal' Zeeman effect is actually observed in the helium spectrum, but in most other spectra the changes produced by the field are more complex; the number of components is greater. However there are indications even in the most complex cases that the simple theory provides an approximation to the truth. For as the strength of the field is increased, the complex series of components tends to coalesce into the simple triplet or doublet of the normal effect—a change discovered by Paschen and Back and generally known by their names. Again, though the separations of the components are not given exactly by (66), the ratio of the separations to that of the normal components is always the ratio of simple integers. (Runge's rule.)

But of course this theory of the Zeeman effect, due to Lorentz, cannot be reconciled with the ideas that we are discussing; it is based on the old conception of radiation emitted from a system in a state of steady vibration, while we must interpret the change produced in a magnetic field as a change in the atomic states between which a transition takes place. The problem might be attacked directly according to the general principles of our theory, by determining what are the mechanical orbits, what are the special oscillating or separable coordinates suitable for their description, and what orbits appear possible when the quantum relations (47) are applied to these coordinates. However another method will be followed here; it can be shown to lead to precisely the same result as the standard method, but the physical meaning of the assumptions involved is much more obvious. And since throughout it is our object to show that an explanation of the facts can be based on assumptions that are simple and natural rather than on assumptions of great formal generality, the alternative method is preferable. Ultimately the justification of any formal principles of complete generality that can be stated is to be found in their reduction in selected cases to principles that are more directly convincing.

Larmor has shown that if an electron is moving in any way under any forces and if its motion is described relative to any axes, then the effect of the establishment of a magnetic field is to change the motion in such a manner that, if the new motion is

referred to axes rotating round the direction of the field with the angular velocity $\Omega = \frac{He}{2\mu c}$ relative to the old axes, then the new motion referred to the new axes is described by precisely the same equations as the old motion referred to the old axes. The effect of the magnetic field is to superimpose on the whole system a rotation with angular velocity Ω. This proposition enables us to deduce very simply how the mechanical states of the atom will be changed by the magnetic field; and, further, it suggests a very obvious assumption concerning the way in which the possible states are changed. It is natural to assume that the possible states in the presence of the magnetic field are those which, when referred to the new axes, have the same equations as the possible states in the absence of the magnetic field referred to the old axes. This assumption is a natural consequence of regarding the change produced by the field as a change of axes, and on it our treatment of the Zeeman effect is founded[1].

Our first step then must be to determine what are the possible states in the absence of the field. It may seem that we have already solved this problem on p. 66, but the mere presence of the field makes a difference which has not been considered; it turns orbits of two degrees of freedom into orbits with three degrees. For so long as there is no field, there is nothing to distinguish one plane in which an orbit may lie from another; as soon as there is a field, its direction enables us to distinguish between different planes, and, though the orbits remain plane, a third coordinate must be introduced to distinguish those which lie in one plane from those which lie in another. The natural way to introduce such a third coordinate is to substitute for the coordinate ϕ, the azimuth in the plane of the orbit, the two angles θ, ψ, which are respectively the latitude and longitude measured from and in an equatorial plane passing through the nucleus and perpendicular to the field. Our three-dimensional coordinates are then the familiar Eulerian

[1] It should be mentioned that the assumption is a particular case of a much more general assumption due to Ehrenfest, and called by him the principle of adiabatic invariance. This principle is of great importance in the general quantum theory applied to all transformations of energy, and will be discussed in the chapter dealing with that theory.

set r, θ, ψ. The kinetic energy of the electron is

$$W_{\text{kin.}} = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\psi}^2 \sin^2\theta) \quad \text{.....(67),}$$

and from (40)

$$p_r = \mu\dot{r}, \quad p_\theta = \mu r^2\dot{\theta}, \quad p_\psi = \mu r^2\dot{\psi}\sin^2\theta \quad \text{.....(68).}$$

We have now three quantum relations similar to (47)

$$\int p_r dr = m_1 h\,; \quad \int p_\theta d\theta = m_2 h\,; \quad \int p_\psi d\psi = m_3 h \quad \text{...(69).}$$

The three partial quantum numbers m_1, m_2, m_3 are simply related to the two numbers of (52), (53), which will now be written m_1', m_2'. For, since r is the same for both sets of coordinates, $m_1 = m_1'$. Again, by comparing the expressions for $W_{\text{kin.}}$ in (59), (67), we see that

$$p_\phi\dot{\phi} = p_\theta\dot{\theta} + p_\psi\dot{\psi} \quad \text{.....(70).}$$

Integrating over a complete period, which must be the same for θ, ϕ and ψ, we have

$$\int p_\phi d\phi = \int p_\theta d\theta + \int p_\psi d\psi \quad \text{.....(71),}$$

or

$$m_2' = m_2 + m_3 \quad \text{.....(72).}$$

Just as before, in passing from a circular to an elliptical orbit, we split up the total quantum number into the sum of two partial quantum numbers, so here we split up the single partial number m_2' into the sum of two others, m_2 and m_3.

Further it can be seen from geometry that p_ψ is the component of p_ϕ resolved in the equatorial plane. If α is the angle between the plane and the plane of the orbit

$$p_\psi = p_\phi \cos\alpha \quad \text{.....(73),}$$

or since p_ϕ, and therefore p_ψ, is constant round any one orbit

$$m_3 = m_2' \cos\alpha \quad \text{.....(74),}$$

and

$$\cos\alpha = \frac{m_3}{m_2 + m_3} \quad \text{.....(75).}$$

Since m_2 and m_3 are integers, (75) implies that only certain values of $\cos\alpha$ are possible, namely such as are the ratios of two integers. Accordingly, the presence of a field imposes a new limitation on the possible orbits of the electron; in addition to the limitations imposed formerly on the dimensions and on the eccentricities of the orbits, there is now also a limitation on their orientation; their planes must make certain definite angles with

the direction of the field. This conclusion is interesting, but it does not seem at present to be capable of an interpretation that can be tested by experiment; and it will not be discussed further. But it may be pointed out that the question arises whether the limitation is removed when the strength of the field becomes infinitesimal, and, if it is not, from what plane this angle α is then to be measured. Difficulties about what happens 'in the limit' are inseparable from theories based on a fundamental conception of discontinuity.

Having now determined the orbits in the absence of the field, we can easily find what changes the field will produce. When the magnetic field is present, the kinetic energy of the possible states, relative to the axes rotating with angular velocity Ω, is again given by (67); the kinetic energy relative to the axes at rest will be given by substituting $\dot{\psi} + \Omega$ for $\dot{\psi}$; i.e.,

$$W'_{\text{kin.}} = \frac{\mu}{2}\{\dot{r}^2 + r^2\dot{\theta}^2 + r^2(\dot{\psi} + \Omega)^2 . \sin^2\theta\} \quad \ldots\ldots\ldots\ldots\ldots\ldots(76)$$

$$= \frac{\mu}{2}\{\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\psi}^2 \sin^2\theta\} + \mu r^2\dot{\psi}\sin^2\theta . \Omega \ldots(77);$$

for, since we know the changes considered are small and proportional to Ω, or to H, we may neglect the term in Ω^2. But from (68)

$$\mu r^2\dot{\psi} . \sin^2\theta = p_\psi,$$

and, since p_ψ is constant round the orbit, we have from (69)

$$p_\psi = \frac{m_3 h}{2\pi} \quad \ldots\ldots\ldots\ldots\ldots\ldots(78).$$

Accordingly if we compare (77) and (67), and write

$$\delta W_{\text{kin.}} = W'_{\text{kin.}} - W_{\text{kin.}}$$

for the change due to the field,

$$\delta W_{\text{kin.}} = p_\psi . \Omega = m_3 h . \frac{He}{4\pi\mu c} \quad \ldots\ldots\ldots\ldots\ldots(79).$$

On the other hand, the potential energy of the possible state with the field will be the same as that without, so that (79) represents the whole difference in energy due to the field, and we may write $\delta W_{\text{kin.}} = \delta W$.

If, then, a line in the spectrum is given out by the passage from a state of energy W_a, in which $m_3 = m_3$, to one of energy W_b,

in which $m_3 = n_3$, the frequency which, in the absence of the field, is given by $W_b - W_a = h\nu$ will be given, in the presence of the field, by

$$W_b + \delta W_b - W_a - \delta W_a = h\nu' \quad \ldots\ldots\ldots\ldots(80).$$

Hence $\delta\nu$, the change in frequency, will be given by

$$\delta\nu = \frac{1}{h}(\delta W_b - \delta W_a) = (n_3 - m_3) \cdot \frac{He}{4\pi\mu c} \quad \ldots\ldots(81).$$

It will be seen that h disappears from the equation and that (81) is identical with (66) if $n_3 - m_3 = 1$. Accordingly our theory predicts the normal Zeeman effect, in so far as frequency is concerned, as one of many possible alternatives; the components of the normal doublet and the displaced components of the normal triplet correspond to transitions between possible states differing by one unit in the partial quantum number m_3, the undisplaced line to a transition in which there is no change in that number. But so far it seems to predict also that lines should occur with integral multiples of the separation of the normal doublet; such lines have never been observed even in the complex cases when the effect is not normal. The explanation of this apparent discrepancy will concern us later. But it is clear that the theory can give no account of the complex cases in which separations are submultiples of the normal; in this matter it is subject to the same limitations as the original theory of Lorentz. It is remarkable that the region in which the older theory was most successful is that in which the newer theory is least successful. The new ideas, though they are not less satisfactory than the old, add nothing so far to our knowledge of the Zeeman effect.

The Stark effect. But with the Stark effect it is otherwise. Many experimenters had sought to detect an influence of an electric field upon spectra, but until 1913 none had succeeded. Their failure arose chiefly from the fact that a gas when emitting its spectrum is always conducting, so that it was impossible to maintain in it the strong electric field necessary to obtain a measurable displacement of the lines. However Stark at length succeeded by applying the field to a beam of canal-rays after they had passed through a fine hole in the kathode and entered a space

in which the density of the gas was reduced by pumping far below that of the gas in which the rays were excited, and in which therefore a strong field could be maintained, unaffected by the discharge. In order to establish such a field all that was necessary was to place parallel to the kathode and at a short distance from it a plate of wire gauze maintained at a suitable difference of potential from the kathode. By such means it is practicable to examine the effect on the spectrum of the canal-rays of fields of the order of 10^4—10^5 volt/cm. Almost at the same time, and independently, Lo Surdo discovered the same effect of an electric field on spectra by making observations on the light emitted from the kathode dark space of a discharge tube, where very great electric fields can be maintained, especially at low pressures. But his method is less well suited to quantitative observations because the intensity of the field is less easily measured and controlled.

Stark's first measurements were made on the Balmer series of hydrogen. He found that each line was split into a number of components, the number increasing with the frequency of the line. The differences of the frequencies of these components from that of the undisplaced line were proportional to the field and were all simple integral multiples of a single frequency difference, the same for all lines in the same field. When the lines were observed perpendicular to the field they were seen to be plane-polarised, some with the electric vector along the field, some in the direction at right angles to it, as in the Zeeman effect. When viewed along the field, only the latter components were visible and were unpolarised. Similar results were found in helium, but in this element the components are more complex.

The complete explanation of the frequency of these components is a direct consequence of Sommerfeld's principle and one of the most striking confirmations of it. But the necessary analysis is complicated and can only be sketched. The problem is to determine how the orbits of the revolving electron of the hydrogen atom will be affected by a uniform electric field. It is found that projections of the perturbed orbits on a plane containing the direction of the field are somewhat similar to those shown in Fig. 3; they lie within envelopes bounded by four lines cutting orthogonally. But these

lines are not straight, as in Fig. 3. If rectangular coordinates are taken with the direction of the field as axis of x (but *not* with the origin at the nucleus), the opposite sides of the envelope are arcs of the parabolas belonging to the two families

$$\frac{y^2}{\xi^2} + 2x = \xi^2; \quad \frac{y^2}{\eta^2} - 2x = \eta^2 \quad \ldots\ldots\ldots\ldots\ldots(82);$$

ξ and η are constants for a single parabola but vary from one curve to another of the same family. One pair of opposite sides of the envelope consists of parts of parabolas with different values of ξ, the other of parabolas with different values of η. Any point within the envelope will be the point of intersection of a pair of parabolas, one from each family, and the values of ξ and η for these parabolas intersecting at the point may be used as its coordinates. And these 'parabolic' coordinates are those which are suitable for our purpose, according to p. 72; for the orbit oscillates regularly from ξ for one side of the envelope to ξ for the other, and from η for one side to η for the other. As the third coordinate may be taken ψ, the azimuth of a plane containing the direction of the field and the line joining the electron to the nucleus.

With such coordinates, the energy of an electron moving in an orbit round a nucleus of charge E in an electric field of intensity F can be shown, by direct algebraic transformation from rectangular coordinates, to be given by

$$W_{\text{kin.}} = \frac{\mu}{2}\{(\xi^2+\eta^2)(\dot{\xi}^2+\dot{\eta}^2) + \xi^2\eta^2\dot{\psi}^2\} \quad \ldots\ldots\ldots(83),$$

$$W_{\text{pot.}} = \frac{1}{2(\xi^2+\eta^2)}\{-4eE + eF(\xi^4-\eta^4)\} \quad \ldots\ldots(84).$$

From (40)

$$p_\xi = \mu(\xi^2+\eta^2)\dot{\xi}, \quad p_\eta = \mu(\xi^2+\eta^2)\dot{\eta}, \quad p_\psi = \mu\xi^2\eta^2\dot{\psi} \quad \ldots(85),$$

so that

$$W = \frac{1}{2\mu(\xi^2+\eta^2)}\left\{p_\xi{}^2 + p_\eta{}^2 + \left(\frac{1}{\xi^2}+\frac{1}{\eta^2}\right)p_\psi{}^2 - 4\mu eE + \mu eF(\xi^4-\eta^4)\right\}$$
$$\ldots\ldots(86).$$

The fundamental quantum relations must now be introduced,

$$\int p_\xi d\xi = m_1 h; \quad \int p_\eta d\eta = m_2 h; \quad \int p_\psi d\psi = m_3 h \quad \ldots\ldots(87).$$

Each integral of (87) has to be taken over the range, within which

the coordinate concerned passes through the whole series of its values and returns to the original value. The calculation of the integrals involves complicated analysis, which it would be useless for those to attempt who are not familiar with the method of the 'separation of the variables'; it will not be attempted here. The result is to show that the only values of the energy from (86) consistent with (87) are given by

$$W = W_0 - \frac{2\pi^2\mu e^2 E^2}{h^2}\cdot\frac{1}{m^2} - 3\frac{h^2F}{8\pi^2\mu E}(m_2 - m_1)\,m$$
$$+ \text{terms in higher powers of } F \quad \ldots\ldots(88),$$

where $m = m_1 + m_2 + m_3$. As before, m is the total quantum number, m_1, m_2, m_3 the partial quantum numbers.

The first two terms on the right are identical with (63) and represent the undisplaced terms of the Balmer series; the change due to the field, δW, is given by the third term; for the remainder may be neglected, since the effect of the field is small. Hence, as when we were dealing with the Zeeman effect, the effect of the field on the frequency of a line which represents the passage from a state of energy W_a, characterised by m_1, m_2, m_3, to a state of energy W_b, characterised by n_1, n_2, n_3, will be given by

$$\delta\nu = \frac{1}{h}(\delta W_b - \delta W_a) = \frac{3hF}{8\pi^2\mu E}\{m(m_2 - m_1) - n(n_2 - n_1)\} \quad \ldots(89).$$

Since the m's and n's are all integers, the values of $\delta\nu$, the displacements of the lines due to the field, will all be integral multiples of the quantity $\frac{3hF}{8\pi^2\mu E} = \Delta$ (say), of which the numerical value for hydrogen ($E = e$) is $\Delta = 6{\cdot}45 \times 10^{-5} F$, where F is in volt/cm., and $\delta\nu$, as usual, in reciprocal centimetres. This representation of all displacements as integral multiples of a minimum displacement agrees with observation and, in hydrogen, the calculated value of Δ agrees with that found experimentally well within the error of the measurements, which arises largely from the uncertainty in the value of F. Again, it is found, as predicted by (89), that the minimum displacement, expressed as a difference in frequency, is the same for all lines in the same spectrum and is independent of the frequency.

Accordingly in all that concerns the minimum displacement, of which all other displacements must be integral multiples, the theory is completely concordant with the facts. We have now to inquire whether the theory predicts rightly what integral multiples of Δ should occur. These integral multiples will be the values of $m(m_2 - m_1) - n(n_2 - n_1)$, and thus depend on both the total and the partial quantum numbers. The question will be best treated by taking an example, the H_α line. The frequency of this line is given by $\nu = N\left(\frac{1}{2^2} - \frac{1}{3^2}\right)$; it represents a transition from any state in which the total quantum number is 3, to one in which it is 2; for it must be remembered that the unperturbed frequencies are determined wholly by the total quantum number and that the partial numbers become important only when perturbations alter the energies of the members of the group of orbits which have all the same total number. The orbits which have the same total number $m = 3$ may have the following values for their partial numbers m_1, m_2, m_3:

(3, 0, 0), (2, 1, 0), (2, 0, 1), (1, 2, 0), (1, 1, 1),
(1, 0, 2), (0, 3, 0), (0, 2, 1), (0, 1, 2), (0, 0, 3);

the orbits for which $n = 2$ may have for their partial numbers:

(2, 0, 0), (1, 1, 0), (1, 0, 1), (0, 2, 0), (0, 1, 1), (0, 0, 2).

Since there are 10 alternatives for $m = 3$ and 6 for $n = 2$, there are 60 possible ways in which the transition in question may occur and 60 ways in which the H_α line may be emitted. But some of the alternatives can be rejected as impossible; we may reject all alternatives for which m_3 or n_3 is zero. The reason is, as on pp. 66, 67, that these orbits pass through the nucleus. As Fig. 3 shows, the projection of the orbit on a plane containing the direction of the field covers in time all the area inside the quadrilateral, which includes the nucleus; so that the electron can only avoid the nucleus by having a rotation round it and a p_ψ different from 0.

The remaining possible transitions are shown in the following table. At the head of each column is a possible orbit for which $m = 3$, and at the left of each row a possible orbit for $n = 2$. Under the symbols for the orbits are the corresponding values of

$m(m_2-m_1)$ and $n(n_2-n_1)$: in a cell of the table are the values of $M = m(m_2-m_1) - n(n_2-n_1)$ which, when multiplied by Δ, are the displacements of the components into which the line is split by the electric field.

TABLE III.

	(2, 0, 1) −6	(1, 1, 1) 0	(1, 0, 2) −3	(0, 2, 1) +6	(0, 1, 2) +3	(0, 0, 3) 0
(1, 0, 1) −2	−4	+2	−1	+8	+5	+2
(0, 1, 1) +2	−8	−2	−5	+4	+1	−2
(0, 0, 2) 0	−6	0	−3	+6	+3	0

We conclude that on each side of the undisplaced line ($M=0$), there will be components separated from it by frequency differences Δ, 2Δ, 3Δ, 4Δ, 5Δ, 6Δ, 8Δ. Of these predicted components those corresponding to $M=0$, 1, 2, 3, 4 have been observed; those corresponding to $M=5$, 6, 8 have not been observed. The theory predicts all observed components, but also some not observed.

If the same calculation were carried out for the H_β line, for which $m=4$, $n=2$, we should find predicted components, symmetrical with respect to the undisplaced line, for which $M=0$, 2, 4, 6, 8, 10, 12, 14; of these all but the last have been observed and no other components are observed. And in general, if we proceeded to other lines, we should find the theory predicting all the lines that are observed, but also some that are not. As the total quantum number increases, so does the number of observed and predicted components. In general the agreement between theory and experiment is very remarkable; the outstanding discrepancy of the predicted but not observed lines will concern us later.

The fine structure. There remains for consideration the 'fine structure' of the hydrogen and helium lines. These lines, even in the absence of magnetic and electric fields, are not strictly single, but are made up of a number of very close components. The theory suggests that the separation of these components must be due to some form of perturbation which distorts the perfect Keplerian orbits and makes the energies of ellipses of different eccentricities not strictly the same. What is the source of this perturbation?

The answer given by Sommerfeld is that there is no real perturbation, but there is an error in our calculations which has the effect of a perturbation. We have assumed throughout that the orbits of the atom are in accordance with Newtonian mechanics. But we know now that Newtonian mechanics is not strictly true, even in its application to those moving systems which it was designed to explain; it is not strictly true for the solar system. The mechanics which we must substitute for it is that of Einstein and Minkowski, the mechanics of relativity. All the considerations urged for the applicability of mechanics of any kind point to the substitution; there can be no reason for using in the quantum theory principles known to be false even when quantum principles have no relevance. Let us therefore revise our calculations in the light of relativity.

Fortunately there is no need to discuss the very abstruse and difficult conceptions that have come to be associated with that word. For our purpose the change introduced by relativity amounts to no more than this, that for a mass of a particle independent of the velocity we must substitute a mass varying with the velocity,

$$\mu = \mu_0 (1 - \beta^2)^{-\frac{1}{2}} \quad \text{.....(90),}$$

where μ_0 is constant and $\beta = v/c$, if v is the velocity of the particle. A particle revolving under a central force still maintains unchanged during its revolution its angular momentum p, but the factor in p which represents the mass must now be considered to depend on the velocity at the moment.

If the central force is that between an electron and a nucleus, the orbit, instead of being a Kepler's ellipse represented by (55), is represented by

$$\frac{a(1-\epsilon^2)}{r} = 1 + \epsilon \cos \gamma\phi \quad \text{.....(91),}$$

where

$$\gamma = 1 - \frac{v_0^2}{c^2} \quad \text{.....(92),}$$

and v_0 may be described here as the mean velocity of the electron in its orbit. The only difference between (55) and (92) lies in the appearance of the factor γ before ϕ. This factor, being nearly but not quite 1, implies that the orbit will no longer be strictly periodic; r will return to its original value not when ϕ changes by 2π, but

when it changes by the slightly greater angle $\frac{2\pi}{\gamma}$. The axis of the ellipse shifts during a revolution by the angle $d\phi = 2\pi . \frac{\gamma - 1}{\gamma}$ and the orbit is like that shown in Fig. 4. This is the irregularity in the motion of the planet Mercury which was for 200 years the chief outstanding difficulty of Newtonian astronomy and which provided the first of the great confirmations of Einstein's doctrine of relativity.

We have then a typical case of a conditionally periodic orbit and can apply directly the principles enunciated for such orbits. The special oscillating and separable coordinates are r, ϕ and the quantum relations are again (52), (53). The integral $\int p_\phi d\phi$ is to be taken as before from 0 to 2π, and since $p_\phi = p = \text{const.}$, we have again $2\pi p = m_2 h$. The integral $\int p_r dr$ is to be taken over a complete oscillation of r, or over the range within which $\gamma\phi$ (and not ϕ) changes from 0 to 2π. It follows that in place of (58) we get

$$1 - \epsilon^2 = \frac{m_2^2 \gamma^2}{(m_1 + m_2\gamma)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(93).$$

To each of the possible ellipses of p. 66 we get a corresponding orbit consisting of an ellipse of very nearly the same eccentricity, of which the axis rotates through the angle $2\pi . \frac{\gamma - 1}{\gamma}$ during each revolution.

W, the energy of the orbit, can now be obtained by straightforward but rather lengthy calculation. The result is

$$\frac{W}{\mu_0 c^2} = 1 - \left\{1 + \frac{\alpha^2}{(m_1 + \sqrt{m_2^2 - \alpha^2})^2}\right\}^{-\frac{1}{2}} \quad \ldots\ldots\ldots(94),$$

where

$$\alpha = \frac{2\pi e E}{hc} = 7{\cdot}3 \times 10^{-3} . \frac{E}{e} \text{ (approx.)} \quad \ldots\ldots\ldots(95).$$

Since α is small compared to 1, we may expand (94) in powers of α and neglect those higher than the fourth. (94) then becomes

$$W = \frac{\mu_0 c^2 \alpha^2}{2} \left\{\frac{1}{m^2} + \frac{\alpha^2}{m^4}\left(\frac{1}{4} + \frac{m_1}{m_2}\right)\right\} \quad \ldots\ldots\ldots\ldots(96),$$

where, as usual, $m = m_1 + m_2$, the total quantum number.

From (27) and (95) we see that $\frac{\mu_0 c^2 \alpha^2}{2h} = N$; the finite mass of the nucleus must be taken into account by using for μ_0 the effective mass of the electron in any numerical calculations. (Cf. p. 50.)

ν, the frequency of a term, is given by $W_0 - W = h\nu$, where W_0 is the value of W for the normal state in which $m = m_2 = 1$; accordingly

$$\nu = N\left[1 + \frac{\alpha^2}{4} - \left\{\frac{1}{m^2} + \frac{\alpha^2}{4m^2} + \frac{\alpha^2}{m^4} \cdot \frac{m_1}{m_2}\right\}\right] \quad \ldots\ldots(97).$$

The constant terms, independent of m, are experimentally insignificant. Of the remaining terms the first is that deduced from Newtonian mechanics: the others represent the 'correction' introduced by relativity mechanics. The second depends only on the total quantum number and is therefore the same for all the components of the fine structure of a line. Its presence merely means that the simple formula (7) for the terms of the hydrogen spectrum is not strictly accurate; but since it is now recognised that the terms are not single but complex, it is clear that that formula cannot be valid for all the components of a term. (7) represents the facts with all the accuracy that is possible so long as the lines are regarded as single, and for the frequency of each is taken that of the centre of the complex; the difference between observed and calculated values is no greater than the breadth of the line. When we have discussed the third term and related each component to the partial quantum number that it represents, we can return and endeavour to connect the components of different lines by means of (97). We shall again find that (97) represents the measurements with full accuracy, but the value we shall obtain for N will differ slightly from that obtained before. However by such means we can hardly confirm (97), for it is much more difficult to fix within a given range of frequency $d\nu$ the frequencies of widely separated lines, than to determine small differences of frequency lying within that range between neighbouring components.

Accordingly we shall only consider the term $\frac{N\alpha^2}{m^4} \cdot \frac{m_1}{m_2}$. Since the components are always separated and there is not, as in the Stark and Zeeman effects, an undisplaced line, experimentally distinguishable, from which we can measure separations, we can

only measure separations between components with different values of this quantity and are concerned only with differences in it with different values of m_1 and m_2. Further, since for the same reasons as on p. 66, the value $m_2 = 0$ may always be rejected, the number of components will be equal to m. Accordingly for $m = 1$, there will be only one component; if we are considering only the separation of components, the smallest value of m that we need consider is 2. When m is 2, the separation between the components is

$$\frac{N\alpha^2}{16}\left(\frac{1}{1}-\frac{0}{2}\right)=\frac{N\alpha^2}{16}.$$

This is the greatest separation that can occur; for the greatest separation corresponding to m is $\frac{N\alpha^2}{m^4}\left(\frac{m-1}{1}-\frac{0}{m}\right)$; and, if m is greater than 2, $\frac{m^4}{m-1}$ is greater than 16. For hydrogen this maximum separation, which we shall write Δ_H (with other suffixes for other elements), is given by

$$\Delta_H=\frac{N_H\alpha^2}{16}=\frac{1{\cdot}097\times10^5\times5{\cdot}31\times10^{-5}}{16}=0{\cdot}365\text{ cm.}^{-1}\quad\ldots(98).$$

Δ_H is less than one-thousandth of the distance between the sodium D-lines. The separations of the components of the term in the spectrum with $m = 3$ are easily seen to be $\Delta_H \cdot \frac{8}{27}$ and $\Delta_H \cdot \frac{8}{81}$; those for $m = 4$, $\Delta_H \cdot \frac{1}{8}$, $\Delta_H \cdot \frac{1}{24}$, $\Delta_H \cdot \frac{1}{48}$; and so on. The separations decrease rapidly as m increases.

Let us now consider, not terms of the spectrum, but lines. The Balmer series consists of lines which represent transitions to a state for which $m = 2$ from states for which m has a series of values increasing from three upwards. In virtue of the occurrence of the term $m = 2$, all the lines should be separated into two groups of components, the corresponding members of these groups being separated by the frequency difference Δ_H. The complexity of these two groups will vary with the line and be determined by m; thus in H_α there will be three members of the group, in H_β four, and so on. The members of these two groups will be much closer to each other than to the members of the other groups; the line will appear as a doublet, each member of which will appear as a line attended by scarcely separated 'satellites.'

The expected structure of the H_α line is shown to scale in Fig. 5; on the same scale of frequency the next member of the series, H_β, would be about 30 miles distant.

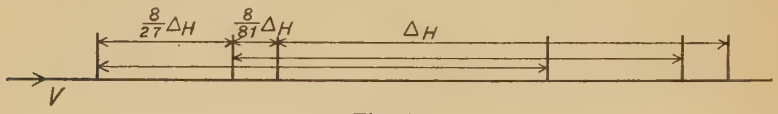

Fig. 5.

Experiment on the whole confirms the theoretical prediction. The lines of the Balmer series (especially the higher members where the satellites are closer) have been shown to be double and the presence of satellites detected. But the separation of the doublet does not accord exactly with expectation; the measured values range from 0·29 to 0·34 cm.$^{-1}$, and undoubtedly tend to be less than the theoretical value 0·365; moreover the separation is apparently not quite the same for different lines. The difference, though small, is outside experimental error; and no thoroughly satisfactory explanation has been offered. But the discrepancy may arise from the fact that when the spectrum is excited the atoms are always subject to an external electric field, and that the resulting Stark effect is comparable with the separation of the components.

But when we pass to spectra other than that of hydrogen, the theory receives brilliant confirmation. It will be seen from (98) that Δ is proportional to N and to α^2; from (27) and (95) both N and α^2 are proportional to E^2, so that Δ is proportional to E^4. This rapid increase of Δ with E is due to the increase of the velocity of the electrons in their orbits which accompanies an increase in the forces due to the nucleus; the difference between the mass in motion and the mass at rest increases very rapidly with the velocity, when it becomes comparable with that of light. Accordingly in spectra arising from orbits round nuclei with more than a single charge, the scale of the separations is greatly increased. For the spark spectrum of helium $E = 2e$, and

$$\Delta_{He} = 16\Delta_H;$$

the opportunities for accurate measurement and detailed testing of the theory are much greater. Paschen has examined with great care the fine structure of the helium lines, and finds almost exact

agreement between theory and experiment. All the components that he could observe were predicted by the theory, though there were some predicted that could not be observed. From the observed separations of the components of the line best suited to accurate measurement he calculated the value of Δ_{He} and thence that of Δ_H. He found

$$\Delta_H = 0{\cdot}3645 \pm 0{\cdot}0045 \text{ cm.}^{-1}.$$

This value agrees perfectly with that predicted by (98), which can be deduced from independent determinations of e, μ_0, h, c.

The structure of X-rays. Perhaps the confirmation is even more striking when the theory is applied to X-rays. We have already seen that X-ray spectra are much more similar to the simple hydrogen and helium spectra than are the optical spectra of other elements. The reason is that the orbits concerned are those in the immediate neighbourhood of the nucleus, the powerful field due to which is scarcely distorted by the presence of other electrons. And X-ray spectra provide peculiarly favourable examples on which to test the part of the theory with which we are now concerned; for, since E is approximately the charge on the nucleus, the ratio E/e will greatly exceed 1 or 2, the values which alone appear in optical spectra. For uranium $E/e = 92$, and the separation between the components should be about 92^4 or seven million times as great as in the hydrogen spectrum. In spite of the lesser accuracy of X-ray measurements, the opportunities for investigation will be greater.

We have concluded that the K-term of the X-ray spectrum represents the orbit for which $m = 1$; the L-term that for which $m = 2$; the M-term that for which $m = 3$. Accordingly the conclusions we have just reached indicate that the K-term should be single, the L-term double, and the M-term triple. Any line or absorption limit in which the K-term is concerned (and not the L or M) should be single; one in which the L-term is concerned double, and so on. The only feature in which the K-term is concerned alone is the K-absorption limit; it is actually found that this limit is single. On the other hand the M-term is difficult to investigate experimentally; the best opportunity of testing the

theory occurs in the L-term. There is definite evidence that this term is double[1]; for there is a frequency difference between two of the absorption limits which appears also between the two components of a K-line (which represents a transition from the L-ring to the K-ring) and between the components of four L-lines (which represent transitions from outer rings to the L-ring). If this frequency difference ($L_{\alpha'} - L_{\beta}$ in (12)) represents the difference between the circle and ellipse of total quantum number 2, which is due to the relativity correction, its value for an element of atomic number Z should be given by

$$\Delta_Z = \Delta_H (Z - z)^4 \quad \ldots\ldots\ldots\ldots\ldots\ldots (99),$$

for we have seen that in the L-ring, the effective charge on the nucleus of an element of atomic number Z is $(Z - z)e$, where z is a constant, the same for all elements.

From the very accurate X-ray measurements that have been made in the last few years, Δ_z can be determined. From the measured values we can deduce z by (99), since we know Δ_H and Z. Actually the calculated z varies somewhat with Z. But it must be remembered that in deducing (97) and (99), we neglected the higher powers of α in the expansion of W; these terms, which are entirely insignificant when $E = e$, become important when E is a large multiple of e. Accordingly a nearer approximation should be made, including higher powers of α. It is unnecessary here to set out the algebra required; it is sufficient to give the final result. For all elements on which measurements are available, lying between $Z = 41$ and $Z = 92$, the calculated value of z lies between 3·44 and 3·97, the mean being 3·6; the differences lie within experimental error.

This value of z is less than that indicated by the line in Fig. 2. But that line, as was noted, is not strictly straight. There should be applied to (35) a correction corresponding to that which distinguishes (97) from (7); when this correction is applied z deduced from (99) agrees with that deduced from (35).

There can be no doubt that the L-doublet is that predicted by (99). Recently attention has been drawn to a second and smaller

[1] It probably has more than two components. For there are three and not two L-absorption limits. The significance of the remaining components is not quite clear, but their existence does not alter the fact that two of the components are related as relativity mechanics predicts.

frequency difference characteristic of the L-lines ($L_\phi - L_\upsilon$ in 12)), and it has been shown that it may probably be attributed to the relativity components of the M-term, which is involved in these lines. Since the M-term corresponds to $m = 3$, the largest separation between its components should be $\Delta_z \cdot \frac{8}{81}$. But z should not be the same for this term as for the L-term, for now there is inside the ring not only the K-ring, but also the L-ring; z should be larger. If the calculations just described for the L-doublet are repeated for this M-doublet, it is found again that the value of z calculated from the measured frequency difference of the doublet is constant for all elements; it is approximately 13. The difference between this z and the z for the L-term should be roughly equal to the number of electrons in the L-ring, which are inside the M-ring but not inside the L-ring. The difference found is $13 - 3{\cdot}6 = 9{\cdot}4$, which is not very different from 8, the suggested number of electrons in the L-ring.

Lines and components. It appears then that some of the different lines of the K-, L-, M-spectra correspond to the different components of the same line in the optical spectrum, the separation being due to the variation of mass with velocity. But all the observed lines have not been accounted for; the remainder are probably due to another source of 'perturbations.' The mutual actions of the many electrons in the atom have been treated so far as merely changing the effective charge on the nucleus; further analysis of their effects is likely to be effected best by the method, familiar in astronomy, of regarding these mutual actions as perturbations of the simple orbits that would be described in their absence. These perturbations must produce a very complex system of terms; for each electron can describe many orbits and its effect on other electrons will depend upon which orbit it is describing. On such grounds we may possibly explain, for example, the discrepancy in equations (13); for the L-orbit concerned in L_α as a final state may differ from that concerned in K_α as an initial state, because the distribution of the remaining electrons is different in the two conditions.

The representation of the many lines of an X-ray spectrum as components of simple lines, separated by perturbations, suggests that

the many lines of a complex optical spectrum may be represented in the same way. If the suggestion is correct, the perturbations involved are probably those due to the mutual action of electrons; these perturbations will also separate terms which have the same total, but different partial quantum numbers. Such considerations have been applied by Sommerfeld to account for the many series of terms in a typical optical spectrum (see p. 34). He supposes that in a term (m, s) the azimuthal quantum number is always 1, in (m, p) always 2, in (m, d) always 3, in (m, b) always 4, the variations in the total quantum number m being due to variations in the radial quantum number. On this view, terms of different series with the same m represent members of the same group of orbits which would have the same energy if the field were due to a single nucleus, but differ in energy when that field is perturbed by the presence of other electrons. The reason why m in (m, p) can never be less than 2 (and so on) is thus immediately apparent.

However no attempt can be made here even to mention all the facts of which the theory suggests some explanation. But before we pass on to another development of Bohr's theory, which is as yet in its infancy, it may be well to insist on the extraordinary success that has attained the working out of its fundamental ideas. The stages through which we have seen them pass are perfectly natural consequences, each of the other; and it is no derogation of the genius of Sommerfeld to say that the developments, leading to such amazing results, which the theory has undergone in his hands were implicit in Bohr's original statement of the connection between homogeneous radiation and possible states, and between possible states and the quantum theory. And none of these results are, to my mind at least, as amazing as those which follow from the substitution of relativity for Newtonian mechanics. A theory which can show that the detailed structure of an X-ray spectrum is a direct consequence of the same principles as the motion of the perihelion of Mercury takes rank at once beside those which related the fall of an apple to the orbit of the moon and the propagation of light to the induction between electric currents. To suspend judgment any longer and to refuse to the conceptions of Bohr, of Einstein, and of Sommerfeld the trust which our predecessors gave to those of Newton and Maxwell is to exhibit, not scientific caution, but a most unscientific inelasticity of mind.

VI. THE INTENSITY OF SPECTRAL LINES

Intensity and probability. Hitherto we have only tried to explain the frequency of the lines in a series spectrum. Spectral lines are characterised by different intensities, as well as by different frequencies; can the theory give any account of these intensities?

The intensity of radiation is measured by the energy emitted in unit time. Bohr's theory predicts definitely how much energy is emitted in each of the elementary processes of which the emission consists; if the frequency is ν, then the energy emitted in each such process is $W = h\nu$. Accordingly it is a direct consequence of the theory that the intensity of the radiation must be the product of the energy emitted when the process occurs by the number of times the process occurs in one second.

This simple and direct consequence of the theory enables some explanation to be offered of why, in a single series, the intensity of the lines usually decreases as the frequency increases. For suppose that the series is excited, as it usually is, by the collisions of electrons with the radiating atoms. In order that a line of frequency ν_1 shall be excited, the atom has to receive at least the energy $h\nu_1$ from the collision with an electron. On the other hand, if it received a greater amount of energy $h\nu_2$, where ν_2 is the frequency of the next line in the series, it would probably emit the line ν_2. Accordingly in order that ν_1 may be emitted, an energy lying between $h\nu_1$ and $h\nu_2$ must be received at a collision. The chance that it shall receive energy within these limits may be expected to decrease with $\nu_2 - \nu_1$; probably also it will decrease as ν_1 and ν_2 increase while $\nu_2 - \nu_1$ is constant; for any theory of collisions that has been proposed so far would lead to such a result. But $\nu_2 - \nu_1$ decreases and ν_1, ν_2 increase as we pass from the earlier members of a series to the later; on both accounts we should expect the chance of an atom receiving at any collision the amount of energy necessary to make it emit one of the lines of a series to decrease very rapidly as the frequency of the line increases. Hence, although the energy emitted in one of

the elementary processes increases with that frequency, we may expect the decrease in the number of times that the process occurs to outweigh that increase, and the intensity of the line to decrease as we pass up the series. (Cf. Chap. X, p. 238.)

In this problem, and in all which concern the relative intensities of different lines in a single spectrum, we have to consider the manner in which the spectrum is excited. Our knowledge of the mechanism of excitation is never sufficient to enable us to make quantitative predictions, and in any case the questions raised do not lie strictly within the limits of our discussion; for we are concerned with the processes characteristic of the atom itself rather than with its reactions with external systems. Accordingly we shall leave this problem on one side—there is really little more known about it—and proceed to another and more limited problem which is more strictly within our province. But before passing to it, we may note that there are some very simple facts which provide strong support for Bohr's theory of spectra, as against the older theories which it has replaced, and might have been quoted on p. 17 as among the reasons why the older theories are untenable. For, according to them, the different lines of the same spectrum, or at least of the same series, are to be regarded as harmonics of the same vibration. Since the form of a vibration is not generally dependent on the manner in which it is excited, the relative intensities of the lines of a single series should be constant and independent of the excitation. Such independence is contradicted by all experiment. The lines of lower frequency in a series are usually stronger than those of higher frequency, but the exact ratio of intensities may be altered considerably. Indeed we have seen (p. 25) that in some conditions it is possible to excite one line without the remainder. The variation in the intensities of the lines is strong evidence for the view that the processes involved in their emission are independent.

In considering the intensities of the lines of the same series we were forced to consider the mechanism of excitation, because the energy associated with the different lines is different, and because the availability of different amounts of energy must depend upon the exciting agency. But if there are two lines, both of which require practically the same amount of energy for their emission, then

the chance that the requisite energy should be available for the emission of one line will be practically equal to the chance that it should be available for the other. In such circumstances, we might expect the relative intensities of the lines to be quite independent of the exciting mechanism; if the intensity of the lines is not the same, the difference must be sought in the internal structure of the atom. If one is more intense than the other and therefore emitted more frequently (for the energy involved in each emission is the same), it must be because the emission of that line, or the transition between the atomic states involved in the emission, is intrinsically more probable than that of the other.

Such circumstances occur when a single line is resolved into a number of very close components. The components which form the fine structure of the hydrogen lines, or those into which lines are split in the Stark and Zeeman effects, differ so little in their frequency and in the energy associated with their emission that the marked differences which are observed in their intensities can scarcely be attributed, like the differences between the intensities of the members of the Balmer series, to differences in the availability of the requisite energy. In all such cases the components correspond to transitions between states of the same total quantum number, but the partial quantum numbers of the states between which the atom passes in the emission of one component are not the same as those between which it passes in the emission of another. The observed differences of intensity suggest that all states of the same total quantum number have not the same intrinsic probability, that the states corresponding to some partial quantum numbers are favoured at the expense of the others, and that the atom is *a priori* more likely to assume one of the more favoured than one of the less favoured states. Of course if such favoured states exist, they probably exist even when the lines are not split into components, and all transitions between the same total quantum numbers give the same line. But in that condition we know at present no method of distinguishing different partial quantum numbers; the experimental evidence on which a theory can be based must necessarily be derived from observations made when the components are separated and the energies associated with the different partial quantum numbers nearly, but not quite, equal.

The Principle of Correspondence. Several principles for determining which of the possible states of an atom, possessing the same total quantum number, are intrinsically more probable have been suggested by examining the observed differences of intensity; but the only one of them which is of really general application is that put forward by Bohr, and that alone we shall consider. Like all principles involved in the new theory, it is not deduced from anything, but suggested as a simple and successful method of accounting for the facts; and also, like all those principles, it is suggested by the consideration of some limiting case. The limiting case from which Bohr's principle arises is that mentioned briefly on p. 56; it arises when the quantum numbers of the possible states considered are all very large compared with the change of quantum number involved in a transition between them. We saw there, in a very simple example, that in this case the predictions of Bohr's theory agreed with those of the older theory (which, for brevity, will be called Maxwell's), so that, since Bohr's theory is certainly true, Maxwell's theory is also applicable.

We must now describe the agreement in rather more general terms; however, for convenience of exposition, we shall not at once proceed to full generality, but shall consider systems of one degree of freedom. According to Fourier's theorem (which depends simply on mathematical conceptions and involves no physical assumptions), the coordinate of any periodic system of one degree of freedom can be expressed as a function of the time by the equation

$$q = \Sigma C_\tau \cos(2\pi\tau\omega + c_\tau) \quad \text{.....(100).}$$

Here ω is the 'fundamental frequency,' τ any positive integer (including 0), C_τ, c_τ coefficients varying with τ; the summation is to be taken for all integral values of τ, but since C_τ may be 0 for some values of τ, those values may not occur explicitly. Now according to Maxwell's theory (which is not purely mathematical, but does involve physical assumptions—although some writers try to disguise them), if q is the position of a charged body moving in the periodic path described by (100), radiation will be emitted of every frequency $\tau\omega$, for which C_τ is not zero. Part of the radiation will have the fundamental frequency ω (unless $C_1 = 0$, an altogether exceptional case), part will have the frequency of the higher harmonics 2ω, 3ω, ..., and, generally, $\tau\omega$. The intensity of any of

these harmonics will be proportional to the square of the acceleration of the charged body in that harmonic; that is to say, the intensity of the τth harmonic will be proportional to $gC_\tau^2(\tau\omega)^4$, where g is the same for all harmonics.

Suppose now that (100) represents the possible orbit of an electron in an atom when the quantum number m is very large. Then other possible states characterised by m', where $m-m'$ is very small compared with m, will differ infinitesimally from the orbit m and may also be described by (100) with unchanged ω, C's and c's. In these conditions, it can be proved quite generally from the Hamiltonian equations that the frequency of the radiation predicted by Bohr's theory corresponding to a transition from the state m to the state m' is $(m-m')\,\omega$, where ω is the fundamental frequency in (100). Now $m-m'$ may have all possible positive integral values, small compared with m, so that to every value of τ will correspond a value of $m-m'$, and to every 'harmonic' predicted by Maxwell's theory will correspond a frequency predicted by Bohr's theory. Consequently, if Bohr's theory is true in this limiting case, Maxwell's must also be true, so far as the predicted frequencies are concerned; for the frequencies predicted by Maxwell's theory are precisely the same as those predicted by Bohr's.

But Maxwell's theory predicts not only the frequency, but also the intensity; the intensity and frequency are inevitably associated. In this limiting case, we may assume that, as it predicts rightly the frequency, it predicts rightly the intensity; for we know that in cases better adapted to experiment (e.g. in experiments with Hertzian waves) correct prediction of intensity is associated with correct prediction of frequency. And what it predicts is that the intensity depends, apart from a universal constant, on (1) the frequency of the harmonic, (2) the coefficient C_τ of the harmonic. According to Bohr's theory, the intensity of a line depends on (1) its frequency and (2) the probability of the transition of which the line is the manifestation. The conclusion is obvious that the coefficient C_τ is a measure of the probability of the transition which gives rise to the line of frequency $(m-m')\,\omega$, where $m-m'=\tau$. Maxwell's theory and all mechanical theories must be merely statistically true, and the quantities which appear as amplitudes must really represent probabilities.

The agreement between Maxwell's and Bohr's theories exists only so long as the quantum number m is very large. Bohr suggests that the feature of the agreement that we have just noted persists when the quantum number is not large, and that it is true generally that the coefficients C_τ in the expression of possible orbits by Fourier's method measure the probabilities of transitions between those orbits. To this proposition he gives the name of the Principle of Correspondence[1].

But the principle needs further explanation. When $m - m'$ is comparable with m, the orbits m and m' differ considerably; ω and the C's will be different for the two orbits, and $\tau\omega$ will not give the frequencies predicted by Bohr's theory. Accordingly we must determine more precisely which of the C's are measures of the intrinsic probability of the transition. The principle asserts that this probability will be measured by some function of C_τ for the m orbit and of C_τ' for the m' orbit, τ for both orbits being equal to $m - m'$; and that this function will be such that it increases with both C_τ and C_τ'. If, for example, the radiation is emitted in a change from an orbit characterised by $m = 3$ to one characterised by $m' = 1$, the probability of the transition will be a function of C_2 for the first orbit and of C_2' for the second. And here it must be noted that, contrary to what was hinted on p. 95, we can be concerned only with the probability of a transition between orbits and not with the probability of the occurrence of the orbits between which the transition takes place. For the C_τ which measures the transition will depend, not only on one of the orbits, but on both. If the transition were from $m = 3$ to $m' = 2$, then the probability would be a function of C_1 for the first orbit and not of C_2.

It has so far proved impossible to define analytically what function of the C_τ's represents the probability; and in most cases

[1] The principle, as stated by Bohr, is rather wider than this. He gives it in a form in which he can use it to calculate the frequencies, as well as the intensities, of the lines emitted in the transitions between the slightly perturbed conditionally periodic orbits which we considered in pp. 69—92. His analysis for deducing the frequencies is thus greatly simplified; he avoids all the apparatus of the separation of the variables and the associated choice of coordinates. But his method, though mathematically simpler, does not appear to me so direct a consequence of the fundamental assumptions of the whole theory; and I have therefore preferred to sketch, however briefly, the methods of Sommerfeld, Epstein and Schwartzchild.

the predictions of the principle are only qualitative. It indicates that if for any particular transition both C_τ and C_τ' are large, the corresponding radiation will be particularly intense; but it cannot say whether the radiation in a transition for which C_τ is large and C_τ' small will be more intense than that in a transition in which both C_τ and C_τ' are intermediate. But in one case, which is of very great importance, it gives an accurate prediction. If there are two possible orbits, m and m', for both of which C_τ is zero for the same value of τ, then the intensity of the radiation emitted in a transition between two such orbits will be zero, if $m - m'$ is equal to τ. This is a clear consequence of the interpretation of the C's as intrinsic probabilities[1].

A simple and striking example of the application of the principle may make the matter clearer. It arises, not in the theory of series spectra, but in the theory of complete radiation (Chap. X) based on Planck's conception of a linear harmonic oscillator. We have seen that the assumption $E = mh\nu$ is a direct consequence of Sommerfeld's principle; but it is necessary to Planck's theory that m should always be unity. This conclusion is a direct consequence of the Principle of Correspondence. For since the oscillator is harmonic, the Fourier series degenerates to a single term, and (100) becomes

$$x = C \cos 2\pi (\omega t + c) \quad \ldots\ldots\ldots\ldots\ldots\ldots (101).$$

That is to say, C_τ is always zero except when $\tau = 1$. The only possible states between which a transition can take place are consequently those in which $m' - m = 1$, the quantum number changes by one unit, and the energy changes by $h\nu$. This is Planck's assumption—for he does not really assume that, in the formula $E = mh\nu$, m is always 1, but merely that an oscillator can only take up or part with one unit of energy.

[1] It is suggested in the text that the probability of a transition between two orbits is determined wholly by the C's for those orbits. But in general there are mechanical states, which are not possible states, intermediate between the possible orbits. Ought we to take into account the C's for these intermediate states? There is some evidence that we ought; and that even if C_τ is zero for both the possible orbits, the radiation emitted in the transition, though very weak, is not strictly zero unless C_τ for the same τ is also zero (as it often is) for all the intermediate orbits that are not possible.

Extension of the Principle. The extension of this principle to systems of n degrees of freedom is simple. Such systems possess as many fundamental frequencies ω as they have degrees of freedom. If the special oscillating or separable coordinates are used, the motion is described by n equations of the form (100), each involving one of these coordinates. With each coordinate q_k is associated a partial quantum number m_k. It can be shown, as before, that when m_k and m_k' are very large compared with the difference between them, the frequency of the radiation emitted in a transition from m_k' to m_k, according to Bohr's theory, is equal to $\tau\omega_k$, where ω_k is the fundamental frequency of q_k and $\tau = m_k' - m_k$.

In order to apply Maxwell's theory it is necessary to express the motion in rectangular coordinates which will not be in general oscillating coordinates. In terms of any but the special coordinates, the motion is given by

$$x_s = \Sigma C_{\tau_1,\tau_2,\ldots\tau_n} \cos 2\pi\{t(\omega_1\tau_1+\omega_2\tau_2+\ldots+\omega_n\tau_n)+c_{\tau_1,\tau_2,\ldots\tau_n}\} \quad (102).$$

Here the ω's are the fundamental frequencies of the oscillating coordinates, the τ's integers, the C's and c's dependent on the τ's; the summation is to be taken over all possible sets of the τ's. In respect of these non-oscillating coordinates, the system has no fundamental frequency, except in the very special case when the ω's are commensurable and have a least common multiple ω_0; in that case (102) reduces to (100). But Maxwell's theory leads to the conclusion that, when the x's are rectangular coordinates, the frequencies of the radiation emitted will be given by

$$(\omega_1\tau_1 + \omega_2\tau_2 + \ldots + \omega_n\tau_n)$$

for all possible sets of the τ's, that the intensities of the radiations will be proportional to $C^2_{\tau_1,\tau_2,\ldots\tau_n}(\omega_1\tau_1 + \omega_2\tau_2 + \ldots + \omega_n\tau_n)^4$, and that the radiations of these intensities will be polarised with their electric vectors along x_s. The Principle of Correspondence then indicates that if, in the equation (102) for any rectangular co-ordinate x_s, $C_{\tau_1,\tau_2,\ldots\tau_n}$ is zero for some particular set of τ's in both the orbits between which a transition takes place (and also in all orbits intermediate between these); and if further each $m_k' - m_k$ is equal to the corresponding τ_k, where m_k' and m_k are the partial quantum numbers corresponding to q_k for these two orbits;

then the intensity of the radiation emitted in this transition, polarised with its electric vector along x_s, will be zero. If C is not actually zero, but is very small for both orbits, then the radiation polarised in this direction emitted in the transition will be correspondingly weak; and generally, as before, the C's for any value of τ_k will be a measure of the probability that a transition occurs for which $m_k' - m_k$ is equal to the value of τ_k.

Here again an example will probably make the matter clearer than any general statement. A very simple case arises in both the Zeeman and Stark effects. It can be shown generally that the vibrations of a particle round an axis of symmetry can be resolved into (1) linear vibrations parallel to the axis with frequencies $(\tau_1\omega_1 + \tau_2\omega_2)$ and (2) circular vibrations (compounded, of course, of equal and perpendicular linear vibrations) round the axis with frequencies $(\tau_1\omega_1 + \tau_2\omega_2 \pm \omega_3)$, where the suffix 3 is associated with rotation round the axis and consequently with the quantum number that we have termed 'equatorial,' the suffixes 1 and 2 with any other coordinates. In other words, C is always zero for (1) unless τ_3 is always zero; and is always zero for (2) unless $\tau_3 = \pm 1$. The vibrations (1) correspond to plane polarised radiation with its vector parallel to the field, and the vibrations (2) correspond to circular polarisation in a plane perpendicular to the field. Consequently the principle of correspondence asserts that the intensity of the radiation polarised parallel to the field will be zero unless $(m_3' - m_3)$ is zero; and that of the radiation polarised circularly round the field will be zero unless $(m_3' - m_3) = \pm 1$. That is to say, all the components polarised parallel to the field will be those that arise from transitions in which m_3 remains unchanged; all those polarised circularly round the field (appearing plane polarised perpendicular to the field when the radiation is received at right angles to the field) will arise from transitions in which m_3 changes to one unit; and there will be no components at all for which m_3 changes by more than one unit. Here we have the explanation of the fact, noted on p. 78, that, though at first sight Bohr's theory would seem to predict Zeeman components with integral multiples of the normal separation, such components do not actually occur; for (81) shows that the normal separation corresponds to $(m_3' - m_3) = 1$; and the

principle of correspondence predicts that $(m_3' - m_3)$ can never be anything but 0 or 1. In the Stark effect, (89) shows that the separation depends on m_3, only because it is involved in m. The simple consequence of the principle does not limit so closely the number of components, because components excluded by it may arise by appropriate changes of the other quantum numbers. But it does predict definitely that components for which m_3 changes by 1 should be polarised parallel to the field, those for which it does not change perpendicular to it. An examination of Table IV will show that the following components, if they occur at all, should be polarised parallel to the field: $M = \mathbf{2}, \mathbf{3}, \mathbf{4}, 8$; and the following, if they occur, perpendicular to it: $M = \mathbf{0}, \mathbf{1}, 5, 6$. Of these lines, those printed in heavy type have actually been observed with the correct polarisation.

But not all the predicted lines are observed. To explain their absence, a more detailed inquiry must be conducted. We must actually determine the coefficients C for the initial, final and intermediate states involved in the transition which is represented by these lines. The necessary mathematics is long and complicated, and cannot lead to any perfectly definite conclusion, because it is unknown what function of the C's is to represent the probability of the transition and the intensity of the line. We shall therefore confine ourselves to a brief indication of the result obtained. In the following table the first column gives the value of M for the component concerned, the second states whether it is polarised parallel or perpendicular to the field, the third gives C' for the initial orbit, the fourth C for the final orbit; C' and C correspond to a set of τ's fixed by the values of the m's and n's for the particular component according to the Principle of Correspondence. In the fifth column is $(C'^2 + C^2)$, which we may take provisionally as a rough estimate of the probability of the transition, since it increases with the arithmetical values of C and C'—their sign is of no importance. In the sixth column is the observed intensity on an arbitrary scale.

It will be seen that there is complete agreement between the sequence of magnitude of the figures in the last two columns. The same kind of agreement between theory and experiment is found in the examination of other lines in the spectra of hydro-

gen and helium; there are one or two discrepancies, chiefly in the appearance of faint components with the right frequency but the wrong polarisation; but in general the concordance is such as to leave no doubt that the basis of the estimation of intensity and polarisation is sound.

TABLE IV.

M	Polarised	C'	C	(C'^2+C^2)	Intensity
2	Parallel	0·46	0	0·21	1
3	„	0·51	0	0·26	1·1
4	„	0·62	0·57	0·71	1·2
8	„	0	0	0	0
0	Perpendicular	1·00	1·00	2·00	2·6
1	„	0·75	0·62	0·95	1
5	„	0	0	0	0
6	„	0·05	0	0·00	0

The application of the Principle of Correspondence to the fine structure represented by the relativity 'correction' is much more complex, because it is impossible in optical spectra to ensure complete freedom from the Stark effect. All spectra are probably excited in an electric field, even if it is only that due to neighbouring charged atoms. In optical spectra the separation of the components of the fine structure is so small that it is masked or confused by the Stark effect arising from very small fields. Kramers has studied the matter in great detail and finds in the observations nothing that conflicts with the Principle of Correspondence and much that confirms it, but his discussion is necessarily too involved for summary here. On the other hand, if we turn to X-ray spectra, where the relativity 'correction' is much greater and the Stark effect inappreciable compared with it, we find that we have detailed knowledge of the fine structure of only one line, the L-doublet; and since this has only two components, according to the theory of p. 89, it offers little opportunity for testing our calculations. But so far as tests can be applied they are satisfactory. The Principle of Correspondence predicts that the component of lower frequency should be the more intense, and it is actually found to be the more intense.

To sum up, the Principle of Correspondence in its present form does not lead to results either as definite or as strictly in accordance with experiment as the two principles characteristic of the earliest form of Bohr's theory. Moreover some of the correct conclusions that can be drawn from it follow also from alternative principles which, though of less generality, are equally plausible at first inspection; but all these principles can be interpreted as special applications of the Principle of Correspondence, which holds the field at present as by far the most satisfactory basis for the estimate of the intensity of spectral lines. But the great importance of the principle is more fundamental; it provides a bridge over the gap between continuous and discontinuous theories of motion. It indicates clearly that the continuous theories, in so far as they are correct at all, are correct because they are applied to statistical groups; they are true in the same sense and in the same way that the Second Law of Thermodynamics is true. Their concepts apply not to the individual elements of the group, but to the group as a whole; they represent probabilities. If this view can be developed and it can be shown how and why the wave-theory of light is a statistical representation of the process of which the Principle of Combination represents the elements, then the last great difficulty which attends the modern theory of light emission will be resolved.

VII. BAND SPECTRA

Regularities in Band Spectra. A typical band spectrum consists of a very large number of lines in close sequence, crowded together towards one end of the sequence to form the 'head' of the band, and gradually increasing in separation towards the other end or 'tail.' There are often several heads associated in the same spectrum, so that the entire spectrum is made up of several bands. Such band spectra are usually, if not always, emission spectra; the best known of them is the 'cyanogen' band, which is now believed to be due to the molecule of nitrogen, N_2.

A numerical law relating the frequencies of the lines making up a single band was propounded by Deslandres, about the time that Balmer propounded the first law of line spectra. He showed that these frequencies could be represented by

$$\nu = A + Bm^2 \quad \text{.....(103),}$$

where m is, as before, the series of successive integers, and A and B are constants characteristic of the band, A being the frequency of the head. But the formula seems never to represent the facts quite accurately, and in particular the lines for small m's, which would provide the most searching test of it, are merged in the head and cannot be distinguished. Better agreement can be obtained by introducing powers of m other than the second, together with their appropriate constants, e.g.

$$\nu = A + Bm + Cm^2 \quad \text{.....(104).}$$

But of course each additional constant makes the fit of the formula better, whether or no it has theoretical justification.

Another type of band, which is becoming of great importance, has been discovered in the last ten years in the infra-red spectra of water vapour and other gaseous compounds, chiefly those of hydrogen; it is usually studied in absorption spectra, but can also appear in emission. Bands of this type have a constant frequency

difference between the lines; the numerical law between the frequencies of the lines in these bands being given approximately by

$$\nu = a \pm bm \quad \text{.....(105).}$$

When such bands are found, there are always at least two of them, one in the very distant infra-red (ν about 100), the other on the borders of the visible region (ν about 10,000). Both of these bands, which are doubtless closely connected, have the same value of the frequency difference b; but in the latter negative as well as positive values of m occur and the band is symmetrical about the frequency a, while in the former only positive values occur, for a and b are so nearly equal that negative values of m would mean negative frequencies.

The rotation of the molecule. If the frequencies of the lines of these bands are to be represented as differences between terms, the terms must involve positive powers of m (e.g. m, m^2, etc.) and not negative powers (e.g. $1/m^2$) as in the terms of line spectra. This difference indicates that the possible states of the system, corresponding to the terms, must be sought elsewhere than among the orbits of charged particles under central forces. Bjerrum in 1912 suggested that they were to be found among the rotations of a rigid body[1]. For it is apparently certain that all band spectra arise from polyatomic molecules, and it is known that such molecules, unlike monatomic molecules or atoms, possess rotational energy as part of their thermal energy. The view therefore is plausible that the emission and absorption of radiant energy in band spectra represent changes in this rotational energy.

If molecules were really rigid bodies, the application of Bohr's principles to the suggestion would be simple. If I is the moment of inertia of the body about its axis of rotation, ω the angular velocity of rotation, the energy of rotation is given by

$$W = \tfrac{1}{2} I\omega^2 \quad \text{.....(106).}$$

[1] At the time of Bjerrum's first suggestion, Bohr's theory had not been propounded; accordingly there was no mention of possible states, but only of resonators in Planck's sense. What we are concerned with now is the interpretation of the suggestion in terms of the ideas of Bohr's theory.

The general quantum relations (47) reduce, as for central orbits, to the condition that the moment of momentum must be an integral multiple of $h/2\pi$, or

$$I\omega = \frac{mh}{2\pi} \quad \text{.....(107).}$$

Consequently the terms of the spectrum should be given by

$$\frac{W}{h} = \frac{h}{8\pi^2 I} m^2 \quad \text{.....(108),}$$

and the frequency of the lines should be given by

$$\nu = \frac{h}{8\pi^2 I} (m_1^2 - m_2^2) \quad \text{.....(109).}$$

(109) is similar to Deslandres' Law (103) in involving m^2, but not identical with it; actualy no band spectrum is known which obeys (109). But molecules are not really rigid bodies. They probably consist of two or more nuclei, in which resides practically all the mass of the molecule, held at finite distances apart by their mutual repulsion, compensated in one particular arrangement by the attraction of each nucleus for the electrons which form the rest of the structure. Since the size of a molecule, as determined by gas theory, is approximately independent of the temperature, the arrangement of the nuclei and the distances between them cannot change rapidly with the energy of rotation; the molecule, therefore, like a rigid body, has a moment of inertia which is nearly, but not quite, independent of the velocity of rotation. But the centrifugal force is sure to change somewhat the arrangement of the nuclei and the moment of inertia, and with them the internal energy w of the atom. The term of the spectrum must include this internal energy as well as the energy of rotation, and the frequency of the lines must depend also upon its changes.

Again, since rotation is a pure harmonic motion, represented by a single term of a Fourier series, the Principle of Correspondence, as applied on p. 99 to the possible states of a linear oscillator, indicates that m should never change by more than one unit. Taking these two considerations into account, we conclude that the lines of the band spectrum should be given by

$$\nu = \frac{w - w'}{h} + \frac{h}{8\pi^2} \left(\frac{m^2}{I} - \frac{(m \pm 1)^2}{I'} \right) \quad \text{.....(110),}$$

where w, w' and I, I' are the values corresponding respectively to m and to $m \pm 1$.

If $dI = I' - I$ is small compared to I and m a large integer, and if we write $\delta w = w' - w$, (110) becomes approximately

$$\nu = c_0 + c_1 m + c_2 m^2 \quad \text{.....(111),}$$

where

$$\left.\begin{aligned} c_0 &= \delta w - \frac{h}{8\pi^2 I'} \\ c_1 &= \pm \frac{h}{4\pi^2 I} \\ c_2 &= \frac{h\delta I}{8\pi^2 I^2} \end{aligned}\right\} \text{.....(112).}$$

It should be observed that, in assuming in (111) that c_0 and c_2 are independent of m, we are assuming that the change in the moment of inertia and in the internal energy, when m changes by one unit, is the same whatever the previous value of m; no justification for the assumption can be offered, except that it seems to lead to approximately the right result.

(111) has been applied to the cyanogen band and shown to fit well with experimental measurements. From the value of c_1, I can be calculated and compared with the estimated moment of inertia of the nitrogen molecule. It is found that

$$I = 1{\cdot}4 \times 10^{-39} \text{ gm. cm.}^{-2};$$

if we take the distance between the nuclei to be 10^{-8} cm., the estimated value is $1{\cdot}1 \times 10^{-39}$, a satisfactory agreement showing that the theory is at least plausible.

(111) can be also employed to interpret the bands of constant frequency difference mentioned on p. 106. Here we have to suppose that δI is very small, and the moment of inertia almost independent of the rotation. The frequency difference b in (105) is to be identified with c_1, and again an estimate of the moment of inertia can be obtained; it is again consistent with our knowledge of molecular dimensions. But there is a further point of interest connected with c_0, which, since c_0 is large compared with c_1, must be nearly equal to δw, and must represent the change in internal energy due to the change in speed of rotation.

Let us suppose that the nuclei are held together by some kind of elastic force, so that, according to mechanical principles,

they would execute vibrations if they were displaced and released. If the elastic force were proportional to the displacement, the vibrations would be harmonic and represented by a single term in a Fourier series; but if, as is more probable, it is not simply proportional, the vibrations would be represented by a Fourier series of several terms and there would be several components, all with frequencies $\tau\omega$ which would be integral multiples of the fundamental frequency. Translating this proposition into the ideas of Bohr's theory by means of the Principle of Correspondence, we find that the transitions between possible states connected with the distance between the nuclei should be such that the changes of energy during the alteration of the distance should all be integral multiples of some constant. We should expect therefore, that if c_0 represents a change of energy arising in this way, there should be other bands in the spectrum with the same values of c_1, but with c_0 replaced by τc_0, where τ is an integer. Such bands are actually found. That in the extreme infra-red, mentioned above, corresponds to $\tau = 0$; other bands have been found with $\tau = 2, 3, 4, \ldots$.

This mere sketch of the theory of band spectra must suffice. The subject is as yet little developed, but enough has been done to show that, by means of Bjerrum's fruitful suggestion, the division between band spectra and line spectra may be broken down, both brought within the range of a single all-embracing theory and both made to contribute in equal degree, the one to our knowledge of the atom, the other to our knowledge of the molecule.

INDEX

Absorption limit 5, 28, 29, 89, 90
Atomic number 36
  radius 48, 57
  states
    and electronic orbits 42–48, 64–68
    normal and ionised 21–28
    and permanent structure 29, 55
    and terms 16

Balmer series 9, 33, 55, 79–83, 88
Bohr's theory
  first assumption 15, 16
  generalisation of 60–63
  and mechanics 57–59, 96–99
  second assumption 44–47

Combination, principle of 11–14, 35, 38, 39
Components (of lines) 1, 68, 69, 91, 92, 95
Coordinates, generalised 60, 61
  oscillating 72
  separable 72, 73
Correspondence, principle of 96–104, 107, 109

Doppler effect 1, 17

Fluorescence 5, 28
Fourier's theorem 39, 40, 96–100
Frequency 6, 7

Homogeneous radiation 2, 15, 18

Ionisation potential 22, 24, 25

Lines, breadth of 1
  and Components 1, 68–92, 95
  fine structure of 83–89, 103

Mechanics
  Hamiltonian 41, 59–62
  and radiation 18–21, 39, 40, 57–59, 94, 96, 97
  relativity 84

Nucleus, effective charge on 52, 53, 90
  mass of 49–51, 86

Orbits, circular 42–44
  conditionally periodic 69–73
  elliptical 43, 59, 64–68
  limiting 56, 57, 96, 97
  relativity 84–86

Perturbations 68, 69, 91, 92
Positive rays 30, 78, 79
Probability and intensity 93–96

Quantum numbers 66–68, 75, 76
Quantum theory (Planck) 15, 25, 46, 63, 64, 99

Radiation (or Resonance) potential 25, 28

Satellites (of lines) 1, 88
Series (in spectra) 9–11, 34, 35
  principal, subsidiary, etc., 34, 35, 92
Spectra (optical), absorption 3, 13, 14, 28–28
  arc and spark 2, 22–24, 56, 57
  band 3, 105–109
  carriers of 29, 30
  excitation of 2, 21–30, 93–95
  flame, *see* arc
  intensity of 93–104
  rotation 106–109
  series in 7–14, 31–36
  terms in 9–11
Spectra (X-ray)
  absorption 5, 13, 14, 28, 29
  details of 89–91
  K-, *L*-, *M*-, 4, 5, 36–39, 54, 55, 89–91
  measurement of 4
  series in 8
Stark effect 69, 78–83, 95, 102–103

Terms (in spectra) 9–11

Valency electrons 23, 55

Wave-length 6, 7
Wave-number 7
Wave-theory (of light) 16–18, 104

Zeeman effect 69, 73–78, 101, 102

CAMBRIDGE: PRINTED BY
J. B. PEACE, M.A.,
AT THE UNIVERSITY PRESS

QC
670
C2
ch.15

FEB 1 1 1977

RAYMOND H. FOGLER LIBRARY
DATE DUE

BOOKS ARE SUBJECT TO
RECALL AFTER TWO WEEKS

UNIV OF MAINE
3 7550 00871561 8

S0-FUV-245